# Bayesian precision matrix estimation for graphical Gaussian models with edge and vertex symmetries

Hélène Massam, Qiong Li and Xin Gao

Department of Mathematics and Statistics, York University, Toronto, Canada

**ABSTRACT**: Graphical Gaussian models with edge and vertex symmetries were introduced by Højsgaard & Lauritzen [2008] who also gave an algorithm to compute the maximum likelihood estimate of the precision matrix for such models. In this paper, we take a Bayesian approach to the estimation of the precision matrix. We consider only those models where the symmetry constraints are imposed on the precision matrix and which thus form a natural exponential family with the precision matrix as the canonical parameter.

We first identify the Diaconis-Ylvisaker conjugate prior for these models and develop a scheme to sample from the prior and posterior distributions. We thus obtain estimates of the posterior mean of the precision matrix.

Second, in order to verify the precision of our estimate, we derive the explicit analytic expression of the expected value of the precision matrix when the graph underlying our model is a tree, a complete graph on three vertices and a decomposable graph on four vertices with various symmetries. In those cases, we compare our estimates with the exact value of the mean of the prior distribution. We also verify the accuracy of our estimates of the posterior mean on simulated data for graphs with up to thirty vertices and various symmetries.

**KEY WORDS**: Conditional independence, symmetries, covariance estimation, trees, Diaconis-Ylvisaker conjugate priors, Metropolis-Hastings.

# 1 Introduction

Given an undirected graph $G = (V, E)$ where $V$ is the set of vertices and $E \subset V \times V$ is the set of undirected edges denoted $(i, j)$, a graphical Gaussian model is a family of Gaussian distribution for $X = (X_v, \ v \in V)$ where the conditional independences between the components of $X$ can be represented by means of a graph as follows:

$$(i, j) \notin E \Rightarrow X_i \perp X_j \mid X_{V \setminus \{i,j\}}.$$

The graphical Gaussian models with edge and vertex symmetries, which we here call the colored graphical Gaussian model, have been introduced in Højsgaard & Lauritzen [2008]. These models are defined as graphical Gaussian models with three different types of symmetry constraints: equality of specified entries of the inverse of the covariance matrix $K$, equality of specified entries of the correlation matrix or equality of specified entries of $K$ generated by a subgroup of the automorphism group of $G$. These models are called respectively RCON, RCOR and RCOP models. In this paper, we consider only RCON models which form a natural exponential family with the precision matrix $K$ as the canonical parameter. The model can be represented by colored graphs, where edges or vertices have the same coloring if the corresponding elements of the precision matrix are equal.

Højsgaard & Lauritzen [2008] proposed an algorithm to compute the maximum likelihood estimates of $K$. However, to the best of our knowledge, there is no work for Bayesian estimates of $K$. Since the RCON model is a natural exponential family, we use the Diaconis & Ylvisaker [1979] (henceforth abbreviated DY) conjugate prior for $K$. This yields a distribution similar to the DY conjugate prior for graphical Gaussian models but with the symmetry constraints mentioned above. We will therefore call this conjugate prior the colored $G$-Wishart.

Our sampling scheme is an adaptation of the independent Metropolis-Hastings (henceforth abbreviated MH) algorithm for the $G$-Wishart proposed by Mitsakakis & al. [2011]. In the case of regular graphical Gaussian models (not colored), the DY conjugate prior for $K$ is the so called $G$-Wishart and there are a number of sampling schemes for this distribution: see Piccioni [2000], Mitsakakis & al. [2011], Dobra & al. [2011], Wang & Li [2012], Lenkoski [2013], the more recent ones being generally more efficient than the preceding ones.

However, by the very nature of these sampling schemes, only Mitsakakis & al. [2011] and Dobra & al. [2011] could be adapted to the colored $G$-Wishart. Moreover, we found that adapting Dobra & al. [2011] leads to significant autocorrelation. The sampling scheme that we propose in Section 3 of this paper is therefore an adaptation to the colored $G$-Wishart of the sampling scheme for the $G$-Wishart given by Mitsakakis & al. [2011].

In order to judge the accuracy of our sample, we need to either know the exact value of the expected value of the precision matrix or we need to proceed by simulations. We will do both. The RCON models for $X = (X_v, v \in V)$ that we consider in this paper are natural exponential families with density of the form

$$f(X; K) \propto \exp\{\langle K, XX^t \rangle - \frac{1}{2} \log |K|\}$$

where $|K|$ is the determinant of the precision matrix and $\langle A, B \rangle = \operatorname{tr} AB$ denotes the inner product of the two symmetric matrices $A$ and $B$. Let $\mathcal{G}$ be the colored version of $G$ to be defined in Section 2. The DY conjugate prior for $K$ is then of the form

$$p(K; \delta, D) = \frac{1}{I_{\mathcal{G}}(\delta, D)} \exp\{-\frac{1}{2}(\langle K, D \rangle - (\delta - 2) \log |K|)\}$$

which is itself an exponential family in $K$ and thus the expected value of $K$ is given by the derivative of $\log I_{\mathcal{G}}(\delta, D)$ with respect to the canonical parameter. The difficulty is, of course, to compute the normalizing constant $I_{\mathcal{G}}(\delta, D)$.

Even for the uncolored $G$-Wishart, until recently, one did not know how to obtain the analytic expression of $I_G(\delta, D)$ unless $G$ was decomposable. Using an iterative method and special functions, Uhler et al. [2014] seem to have solved this very difficult problem but their results do not extend to our colored $G$-Wishart with symmetry constraints which, as we shall see, add another level of difficulty. We do give, in Section 4, the explicit analytic expression of the normalizing constant of the coloured $G$-Wishart for some special graphs $\mathcal{G}$: general trees, star graphs, a complete graph on 3 vertices and a simple decomposable model on 4 vertices with various symmetry constraints. For these particular coloured graphs, the analytic expression of $I_{\mathcal{G}}(\delta, D)$ allows us to verify the accuracy of the estimate of the prior mean obtained with the sampling scheme developed in Section 3 by comparing it to the true prior mean obtained by differentiation of $I_{\mathcal{G}}(\delta, D)$. The numerical results are given in Section 5. In Section 6, we compute the posterior mean for two relatively high-dimensional examples

where the underlying graphs are cycles of length 20 and 30. Since we cannot compute the analytic expression of the normalizing constant in those cases, we compare the posterior mean estimate obtained through our sampling method of Section 3 with the true value of $K$ used for simulating the initial Gaussian data. Our numerical results show good accuracy and small relative errors that decrease with sample size, as expected.

## 2    Preliminaries

We will now recall some definitions and concepts that we will need in the sequel. Let $G = (V, E)$ be an undirected graph as defined in the introduction. Let $P_G$ be the cone of $p \times p$ positive definite matrices $X$ with entry $X_{ij} = 0$ whenever $(i, j) \notin E$. It is well-known (see Lauritzen [1996]) that the graphical Gaussian model Markov with respect to $G$ is the set of Gaussian $N(0, \Sigma)$ distributions

$$\mathcal{N}_G = \{N(0, \Sigma) \mid K = \Sigma^{-1} \in P_G\}. \tag{1}$$

The Diaconis-Ylvisaker conjugate prior for the parameter $K$ is the so-called $G$-Wishart distribution (see Roverato [2002]) defined on $P_G$ and with density

$$p(K|\delta, D) = \frac{1}{I_G(\delta, D)} |K|^{(\delta-2)/2} \exp\{-\frac{1}{2}\langle K, D\rangle\},$$

where $\delta > 0$ and $D$, a symmetric positive definite $p \times p$ matrix, are the hyper parameters of the prior distribution on $K$ and $I_G(\delta, D)$ is the normalizing constant, namely,

$$I_G(\delta, D) = \int_{P_G} |K|^{(\delta-2)/2} \exp\{-\frac{1}{2}\langle K, D\rangle\}dK.$$

Let us now define the RCON model. Let $\mathcal{V} = \{V_1, \ldots, V_k\}$ form a partition of $V = \{1, \ldots, p\}$ and let $\mathcal{E} = \{E_1, \ldots, E_l\}$ form a partition of the edge set $E$. If all the vertices belonging to an element $V_i$ of $\mathcal{V}$ have the same colour, we say that $\mathcal{V} = \{V_1, \ldots, V_k\}$ is a colouring of $V$. Similarly if all the edges belonging to an element $E_i$ of $\mathcal{E}$ have the same colour, we say that $\mathcal{E}$ is a colouring of the edges of $G$ and that $(\mathcal{V}, \mathcal{E})$ is a coloured graph.

Consider model (1). If, for $K \in P_G$, we impose the further restrictions that if

$(C_1):$  $m$ is a vertex class in $\mathcal{V}$, then for all $i \in m$, $K_{ii}$ are equal,

$(C_2):$  $s$ is an edge class in $\mathcal{E}$, then for all $(i, j) \in s$, the entries $K_{ij}$ of the precision matrix are

equal, then model (1) becomes a coloured graphical Gaussian model called the $\mathrm{RCON}(\mathcal{V}, \mathcal{E})$ model.

For the computation of the analytic expression of $I_G(\delta, D)$, we will need two special functions, the Bessel function of the third kind and the hypergeometric function $_pF_q$. The Bessel function of the third kind is defined as

$$K_\lambda(z) = \int_0^\infty u^{2\lambda-1} e^{-\frac{z}{2}(\frac{1}{u^2}+u^2)} du.$$

For some special values of $\lambda$, the Bessel function can be given explicitly

$$K_{1/2}(z) = \sqrt{\frac{\pi}{2}} z^{-1/2} e^{-z}, \quad K_{3/2}(z) = \sqrt{\frac{\pi}{2}} (z^{-1/2} + z^{-3/2}) e^{-z},$$

$$K_{5/2}(z) = \sqrt{\frac{\pi}{2}} (z^{-1/2} + 3z^{-3/2} + 3z^{-5/2}) e^{-z}.$$

We will also use the classical formula

$$(\frac{p}{q})^{\frac{\lambda}{2}} K_\lambda(\sqrt{pq}) = \int_0^\infty u^{2\lambda-1} e^{-\frac{1}{2}(\frac{p}{u^2}+qu^2)} du.$$

The hypergeometric function $_pF_q$ is defined by the power series:

$$_pF_q(a_1, \ldots, a_p; b_1, \ldots, b_q; z) = \sum_{k=0}^\infty \frac{(a_1)_k \cdots (a_p)_k}{(b_1)_k \cdots (b_p)_k} \frac{z^k}{k!}$$

where

$$(a)_k = \begin{cases} 1 & \text{if } n = 0 \\ a(a+1) \cdots (a+n-1) & \text{if } n > 0 \,. \end{cases}$$

The derivative of the hypergeometric function $_pF_q(a_1, \ldots, a_p; b_1, \ldots, b_q; z)$ is given by

$$\frac{d}{dz}[_pF_q(a_1, \ldots, a_p; b_1, \ldots, b_q; z)] = \frac{a_1 \cdots a_p}{b_1 \cdots b_q} (_pF_q(a_1+1, \ldots, a_p+1; b_1+1, \ldots, b_q+1; z)) \,. \quad (2)$$

## 3  The coloured $G$-Wishart distribution: a sampling method

### 3.1  The coloured $G$-Wishart

For $G$ an undirected graph, let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ denote its coloured version as defined in Section 2, and let $P_{\mathcal{G}}$ denote the cone of $p \times p$ positive definite matrices in $P_G$ which also

obey the symmetry constraints of $\mathcal{G}$, i.e.

$$P_{\mathcal{G}} = \{K \in P_G \mid (C_1) \text{ and } (C_2) \text{ are satisfied}\} .$$

We define the $CG$-Wishart, i.e. the colored $G$-Wishart, to be the DY-conjugate prior for the parameter $K$ of the RCON$(\mathcal{V}, \mathcal{E})$ model. Its density is

$$p(K|\delta, D) = \frac{1}{I_{\mathcal{G}}(\delta, D)} |K|^{(\delta-2)/2} \exp\{-\frac{1}{2} \text{ tr } (KD)\} \mathbf{1}_{P_{\mathcal{G}}}(K) \tag{3}$$

where $\delta > 0$ and $D$, a symmetric $p \times p$ matrix, are hyper parameters and $I_{\mathcal{G}}(\delta, D)$ is the normalizing constant, namely,

$$I_{\mathcal{G}}(\delta, D) = \int_{P_{\mathcal{G}}} |K|^{(\delta-2)/2} \exp\{-\frac{1}{2} tr(KD)\} dK. \tag{4}$$

We will see that $I_{\mathcal{G}}(\delta, D)$ is finite only for $D$ in the dual cone of $P_{\mathcal{G}}$ which has to be determined for each $\mathcal{G}$. We will derive the dual cones in the special cases that we consider in Section 4.

In this section, following what has been done in Mitsakakis & al. [2011], we want to derive a MH algorithm to sample from the $CG$-Wishart. But in order to do so, following Atay-Kayis & Massam [2005], we want to express the density of the $CG$-Wishart in terms of the Cholesky components of $K$ scaled by $D$. To this end, we consider the Cholesky decomposition of $D^{-1}$ and $K$ written as

$$D^{-1} = Q^T Q, \quad K = \Phi^T \Phi$$

with $Q = (Q_{ij})_{1 \leq i \leq j \leq p}$ and $\Phi = (\Phi_{ij})_{1 \leq i \leq j \leq p}$ upper triangular matrixes with real positive diagonal entries and we will use the variable

$$\Psi = \Phi Q^{-1}.$$

We adapt the MH algorithm given in Mitsakakis & al. [2011] for the $G$-Wishart to the $CG$-Wishart, using the variable $\Psi$ rather than $K$. To do so, we first define

$$v_u(G) = \min\{(i, j) : i \leq j, \ (i, j) \in u \in \mathcal{V} \cup \mathcal{E}\}$$

where the minimum is defined according to the lexicographical order and

$$v(G) = \bigcup_{u \in \mathcal{V} \cup \mathcal{E}} v_u(G).$$

We will write $K^{v(G)} = (K_{ij}| \ (i,j) \in v(G))$ for the free elements of $K$. The zero and colouring constraints on the elements of $K$ determine the free entries $\Phi^{v(G)} = \{\Phi_{ij} : (i,j) \in v(G)\}$ and $\Psi^{v(G)} = \{\Psi_{ij} : (i,j) \in v(G)\}$ of the matrices $\Phi$ and $\Psi$ respectively. Each non-free element $\Phi_{ij}$ and $\Psi_{ij}$ with $(i,j) \notin v(G)$ is a function of the free elements $\Phi^{v(G)}$ and $\Psi^{v(G)}$ that precede it in the lexicographical order. The following two propositions give the expression of the non-free entries in function of the free ones and the free entries of $K$. The first part of each proposition can be found in Roverato [2002].

**Proposition 3.1** *Let $K = \Phi^T \Phi$ be an element of $P_{\mathcal{G}}$. Then the entries $\Phi_{ij}$ are such that*

$$\Phi_{ij} = \frac{K_{ij} - \sum\limits_{k=1}^{i-1} \Phi_{ki}\Phi_{kj}}{\Phi_{ii}}, \qquad\qquad\qquad\qquad\qquad\qquad for \ (i,j) \in v(G)$$

$$\Phi_{1k} = 0, \qquad\qquad\qquad\qquad\qquad\qquad\qquad for \ K_{1k} = 0, \ k = 2, \ldots, p$$

$$(5)$$

$$\Phi_{ij} = -\frac{\sum\limits_{k=1}^{i-1} \Phi_{ki}\Phi_{kj}}{\Phi_{ii}} \qquad\qquad\qquad\qquad\qquad\qquad for \ K_{ij} = 0, \ldots, p, \ i \neq 1$$

$$\Phi_{ij} = \frac{\Phi_{i_u j_u}\Phi_{i_u i_u} + \sum\limits_{k=1}^{i_u-1} \Phi_{ki_u}\Phi_{kj_u} - \sum\limits_{k=1}^{i-1} \Phi_{ki}\Phi_{kj}}{\Phi_{ii}} \qquad for \ K_{ij} \neq 0, \ (i,j) \in u \in \mathcal{V} \cup \mathcal{E}, \ (i,j) \notin v(G)$$

$$(6)$$

$$\Phi_{ii} = |\Phi_{i_u i_u}^2 + \sum\limits_{k=1}^{i_u-1} \Phi_{ki_u}^2 - \sum\limits_{k=1}^{i-1} \Phi_{ki}^2|^{\frac{1}{2}}, \qquad\qquad\qquad\qquad for \ i = 1, \ldots, p.$$

$$(7)$$

*where $(i_u, j_u) = \min\{(i,j) : i \leq j \ and \ (i,j) \in u \in \mathcal{V} \cup \mathcal{E}\}$ in the lexicographical order.*

**Proof.** The first three equations can be found in Roverato [2002]. We will only prove (6) since (7) will follow immediately from it. For all $(i,j) \in u \in \mathcal{V}\cup\mathcal{E}$ and $(i,j) \neq (i_u, j_u) \in u$, by (5), we have that $K_{i_u j_u} = \sum\limits_{k=1}^{i_u} \Phi_{ki_u}\Phi_{kj_u}$ and in general $K_{ij} = \sum\limits_{k=1}^{i} \Phi_{ki}\Phi_{kj}$. Since $K_{ij} = K_{i_u j_u}$, it follows that

$$\Phi_{i_u j_u}\Phi_{i_u i_u} + \sum\limits_{k=1}^{i_u-1} \Phi_{ki_u}\Phi_{kj_u} = \Phi_{ii}\Phi_{ij} + \sum\limits_{k=1}^{i-1} \Phi_{ki}\Phi_{kj}.$$

Equations (6) and (7) follow then immediately. ∎

**Proposition 3.2** *For* $K = Q^T(\Psi^T\Psi)Q \in P_{\mathcal{G}}$ *with* $\Psi$ *and* $Q$ *as defined above, the entries* $\Psi_{ij}$ *of* $\Psi$ *are as follows:*

$$\Psi_{rs} = \sum_{j=r}^{s-1} -\Psi_{rj}\frac{Q_{js}}{Q_{ss}} + \frac{\Phi_{rs}}{Q_{ss}} \quad \textit{for } (r,s) \in v(G),\ r \neq s\,, \tag{8}$$

$$\Psi_{ss} = \frac{\Phi_{ss}}{Q_{ss}} \quad \textit{for } (r,s) \in v(G),\ r = s\,,$$

$$\Psi_{rs} = \sum_{j=r}^{s-1} -\Psi_{rj}\frac{Q_{js}}{Q_{ss}} - \sum_{i=1}^{r-1}\left(\frac{\Psi_{ir} + \sum\limits_{j=i}^{r-1}\Psi_{ij}\frac{Q_{jr}}{Q_{rr}}}{\Psi_{rr}}\right)\left(\Psi_{is} + \sum_{j=i}^{s-1}\Psi_{ij}\frac{Q_{js}}{Q_{ss}}\right) \quad \textit{for } K_{rs} = 0,\ r \neq 1\,,$$

$$\Psi_{1s} = \sum_{j=1}^{s-1}\left(-\Psi_{1j}\frac{Q_{js}}{Q_{ss}}\right) \quad \textit{for } K_{1s} = 0\,,$$

$$\Psi_{rs} = \frac{\Phi_{i_u j_u}\Phi_{i_u i_u} + \sum\limits_{k=1}^{i_u-1}\Phi_{ki_u}\Phi_{kj_u} - \sum\limits_{k=1}^{r-1}\Phi_{kr}\Phi_{ks}}{\Phi_{rr}Q_{ss}} - \sum_{j=r}^{s-1}\Psi_{rj}\frac{Q_{js}}{Q_{ss}} \quad \textit{for } K_{rs} \neq 0, \tag{9}$$

$$(r,s) \in u \in \mathcal{V} \cup \mathcal{E},\ (r,s) \notin v(G)\,,$$

$$\Psi_{ss} = \frac{\left|\Phi_{i_u i_u}^2 + \sum\limits_{k=1}^{i_u-1}\Phi_{ki_u}^2 - \sum\limits_{k=1}^{r-1}\Phi_{ks}^2\right|^{\frac{1}{2}}}{Q_{ss}} \quad \textit{for } s = 1,\ldots,p\,. \tag{10}$$

**Proof.** We will prove (9) and therefore (10). Since $\Phi = \Psi Q$, for $r \neq s$, we have

$$\Phi_{rs} = \Psi_{rs}Q_{ss} + \sum_{j=r}^{s-1}\Psi_{rj}Q_{js}.$$

On the other hand, by (8), we have

$$\Phi_{rs} = \frac{\Phi_{i_u j_u}\Phi_{i_u i_u} + \sum\limits_{k=1}^{i_u-1}\Phi_{ki_u}\Phi_{kj_u} - \sum\limits_{k=1}^{r-1}\Phi_{kr}\Phi_{ks}}{\Phi_{rr}}.$$

It then follows that

$$\Psi_{rs}Q_{ss} + \sum_{j=r}^{s-1}\Psi_{rj}Q_{js} = \frac{\Phi_{i_u j_u}\Phi_{i_u i_u} + \sum\limits_{k=1}^{i_u-1}\Phi_{ki_u}\Phi_{kj_u} - \sum\limits_{k=1}^{r-1}\Phi_{kr}\Phi_{ks}}{\Phi_{rr}}$$

which implies (9) and (10). ∎

Next we compute the Jacobian of the change of variable from $K$ to $\psi^{v(G)}$ in two steps.

8

**Lemma 3.1** Let $K$ be in $P_{\mathcal{G}}$. Let $v_i^G$ be the number $j \in \{i, \ldots, p\}$ such that $(i, j) \notin v(G)$. Then the Jacobian of the change of variable $K^{v(G)} \rightarrow \Phi^{v(G)}$ as defined above is

$$\det(J(K^{v(G)} \rightarrow \Phi^{v(G)})) = 2^{|\mathcal{V}|} \prod_{i=1}^{p} \Phi_{ii}^{p-i+1-v_i^G}$$

where $|\mathcal{V}|$ is the number of vertex color class of $\mathcal{G}$.

**Proof.** Order the elements of both matrices $K$ and $\Phi$ according to the lexicographic order. For $(i, j) \in v(G)$, differentiating (5) yields

$$\frac{\partial K_{ii}}{\partial \Phi_{ii}} = 2\Phi_{ii} \,, \qquad\qquad \frac{\partial K_{ii}}{\partial \Phi_{ks}} = 0 \text{ for } (k, s) > (i, i) \,, \qquad (11)$$

$$\frac{\partial K_{ij}}{\partial \Phi_{ij}} = \Phi_{ii} \,, \qquad\qquad \frac{\partial K_{ij}}{\partial \Phi_{ks}} = 0 \text{ for } (k, s) > (i, j), \ \ i \neq j. \qquad (12)$$

Therefore, the Jacobian is an upper-triangular matrix and its determinant is the product of the diagonal elements. The lemma then follows immediately from the fact that for $i \in \{1, \ldots, p\}$ given, the cardinality of the set $\{(i, j) \in v(G), \ (i, j) \geq (i, i)\}$ is $p - i + 1 - v_i^G$. ∎

**Lemma 3.2** Let $K$ be in $P_{\mathcal{G}}$. Let $d_i^G = | \{j : j \leq i, \ (j, i) \notin v(G)\} |$. The Jacobian of the change of variable $\Phi^{v(G)} \rightarrow \Psi^{v(G)}$ where $\Phi$ and $\Psi$ are as defined above is

$$\det(J(\Phi^{v(G)} \rightarrow \Psi^{v(G)})) = \prod_{i=1}^{p} Q_{ii}^{i-d_i^G}.$$

**Proof.** Order the elements of both matrices $\Phi$ and $\Psi$ according to the lexicographic order. For $(r, s) \in v(G)$, differentiating (8), we obtain

$$\frac{\partial \Phi_{rs}}{\partial \Psi_{ss}} = Q_{ss} \,, \qquad\qquad \frac{\partial \Phi_{rs}}{\partial \Psi_{ij}} = 0 \text{ for } (i, j) > (r, s). \qquad (13)$$

The Jacobian is thus an upper-triangular matrix and its determinant is the product of the diagonal elements. The lemma follows from the definition of $d_i^G$. ∎

**Theorem 3.1** Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be an arbitrary $p$-dimensional colored graph. Then the density of the CG-Wishart distribution expressed in terms of $\Psi^{v(G)}$ is

$$p(\Psi^{v(G)}|\delta, D) = \frac{2^{|\mathcal{V}|}}{I_G(\delta, D)} \prod_{i=1}^{p} Q_{ii}^{p-v_i^G-d_i^G+\delta-1} \prod_{i=1}^{p} \Psi_{ii}^{p-i-v_i^G+\delta-1} e^{-\frac{1}{2} \sum_{i=1}^{p} \sum_{j=i}^{p} \Psi_{ij}^2}. \qquad (14)$$

**Proof.** The expression of $p(\Phi^{v(G)}|\delta, D)$ above follows immediately from the fact that $|K| = \prod_{i=1}^{p} \Phi_{ii}^2$, that $\langle K, D \rangle = \sum_{i=1}^{p} \sum_{j=1}^{p} \Psi_{ij}^2$ and from the expressions of the Jacobians given in Lemmas 3.1 and 3.2. ∎

## 3.2   The sampling algorithm

We now briefly describe the MH algorithm we use to sample from the density (14). We first note that if we make the further change of variables

$$\left(\Psi_{ii}, (i,i) \in v(G), \Psi_{ij}, (i,j) \in v(G), i \neq j\right) \mapsto \left(t_{ii} = \Psi_{ii}^2, (i,i) \in v(G), \Psi_{ij}, (i,j) \in v(G), i \neq j\right),$$

we obtain

$$p(t_{ii}, (i,i) \in v(G), \Psi_{ij}, (i,j) \in v(G), i \neq j | \delta, D) \quad \propto \quad \prod_{i=1}^{p} t_{ii}^{\frac{p-i-v_i^G+\delta}{2}-1} e^{-\frac{1}{2}\sum_{i=1}^{p} t_{ii}} e^{-\frac{1}{2}\sum_{i=1}^{p}\sum_{j=i+1}^{p} \Psi_{ij}^2}$$

and we observe that $t_{ii}^{\frac{p-i-v_i^G+\delta}{2}-1} e^{-\frac{1}{2}\sum_{i=1}^{p} t_{ii}}$ has the form of a $\chi^2_{p-i-v_i^G+\delta}$ distribution.

We denote by $\Psi^{[s]}$ and $\Psi^{[s+1]}$ the current state of the chain and the next state of the chain, respectively. We denote $\Psi'$ the candidate for $\Psi^{[s+1]}$. We also use the notation

$$\Psi_{v(G)^c} = \left(\Psi_{ij}, (i,j) \in v(G)^c\right)$$

where $v(G)^c$ is the complement of $v(G)$ in $V \times V$. For $(i,j) \in v(G)$, an element $\Psi_{ij}^{[s]}$ is updated by sampling a value $\Psi'_{ij}$ from a normal distribution with zero mean and standard deviation equal to one. For $(i,i) \in v(G)$, a element $\Psi_{ii}^{[s]}$ is updated by sampling a value $(\Psi_{ii}^2)'$ from a chi-square distribution with $p - i - v_i^G + \delta$ degrees of freedom. The non-free elements of $\Psi'$ are uniquely defined by the functions in Proposition 3.1 and Proposition 3.2. The Markov chain moves to $\Psi'$ with probability

$$\min\{\frac{h[(\Psi')_{v(G)^c}]}{h[(\Psi^{[s]})_{v(G)^c}]}, 1\},$$

where

$$h(\Psi_{v(G)^c}) = \prod_{(i,i)\in v(G)^c} \Psi_{ii}^{p-i-v_i^G+\delta-1} \exp(-\frac{1}{2}\sum_{(i,j)\in v(G)^c} \Psi_{ij}^2).$$

Finally, we can obtain $K^{[s]} = Q^T(\Psi^{[s]})^T\Psi^{[s]}Q$. Since $h(\Psi_{v(G)^c})$ is uniformly bounded by 1, the chain is uniformly ergodic (the strongest convergence rate in use, see Mengersen & Tweedie [1996]).

We now have a method to sample values of $K$ from the $CG$-distribution, whether it is as a prior or a posterior distribution and thus obtain an estimate of the posterior mean of

$K$. In our MH algorithm, the candidates are drawn independently of the current samples through the proposal density. Thus, the algorithm gives an independence MH chain. Our simulation results in Section 5 will show that the chain has good mixing, low autocorrelation and high proximity to the true $CG$-distribution.

The sample mean will converge to the expected value of $K$. In order to verify the accuracy of our sampling algorithm, we therefore would like to have the exact value of the expected value of $K$ under the $CG$-Wishart. This is done in the next section for some special coloured graphs.

## 4   The exact expected value of $K$ in some special cases

### 4.1   The mean of the $CG$-Wishart

For a given $\mathcal{G}$, the $CG$-Wishart as defined in (3) and (4) clearly form a natural exponential family of the type

$$f(K; \theta)dK = \exp\{\langle K, \theta \rangle - k(\theta)\}\mu(dK)$$

with generating measure $\mu(dK) = |K|^{(\delta-2)/2}\mathbf{1}_{P_\mathcal{G}}(K)$, $\theta = -\frac{1}{2}D$ and cumulant generating function $k(\delta, D) = \log I_\mathcal{G}(\delta, D)$. To verify the accuracy of the sampling method given in Section 3, we will compare the expected value of $K$ under the $CG$-Wishart and the sample mean obtained from a number of iterations of our MH algorithm. From the theory of natural exponential family, we know that the mean of the $CG$-Wishart is

$$E(K) = \frac{\partial\, k(\delta, D)}{\partial\, (-\frac{1}{2}D)} = -2\frac{\partial\, k(\delta, D)}{\partial D}.$$

We therefore need to determine for which values of $\delta$ and $D$ the quantity $I_\mathcal{G}(\delta, D)$ is finite and then compute the analytic expression of $I_\mathcal{G}(\delta, D)$. We need also to differentiate this expression.

We cannot do this in general but we will now consider several particular coloured graphs for which we can compute $I_\mathcal{G}(\delta, D)$. For the corresponding RCON models, we will see that when $\delta > 0$ (except in the case of the star graph with all leaves in the same colour class where we must have $\delta \geq 1$), the normalizing constant $I_\mathcal{G}(\delta, D)$ is finite when $D$ belongs to

11

the dual $P_{\mathcal{G}}^*$ of $P_{\mathcal{G}}$. For any open convex cone $C$ in $R^n$, the dual of $C$ is defined as

$$C^* = \{y \in R^n \mid \langle x, y \rangle > 0, \ \forall x \in \bar{C} \setminus \{0\}\}$$

where $\bar{C}$ denotes the closure of $C$.

In the remainder of this section, for each RCON model, we determine $P_{\mathcal{G}}^*$ and the value of $I_{\mathcal{G}}(\delta, D)$. This will allow us, in Section 5, to verify the accuracy of our sampling method.

All proofs for Section 4 are given in Appendix 1 in the Supplementary file.

## 4.2 Trees with vertices of different colours and edges of the same colour

Let $V = \{1, \ldots, p\}$. Let $T = (V, E)$ be a tree with vertices of different colours and edges of the same colour. An example of such $\mathcal{G}$ is given in Figure 1(a). Let $a = (a_i, i = 1, \ldots, p)^t$ where $a_i \geq 0$ and $b \in R$. Let $S$ be the space of symmetric $p \times p$ matrices. We define the mapping

$$m : (a, b) \in R^{p+1} \mapsto m(a, b) \in S \tag{15}$$

with $m(a, b)$ satisfying the conditions

$$[m(a,b)]_{ii} = a_i, \ [m(a,b)]_{ij} = b = [m(a,b)]_{ji} \text{ for } (i,j) \in E, \ [m(a,b)]_{ij} = 0 \text{ for } (i,j) \notin E.$$

Let $M(\mathcal{G})$ be the linear space of matrices $m(a, b)$ for $(a, b) \in R^{p+1}$. Let $P$ be the cone of $p \times p$ symmetric positive definite matrices. Then

$$P_{\mathcal{G}} = M(\mathcal{G}) \cap P . \tag{16}$$

**Proposition 4.1** *Let $T$ be a tree as described above. The dual cone $P_{\mathcal{G}}^*$ is*

$$P_{\mathcal{G}}^* = \{m(a', b') \in M(\mathcal{G}) \mid a' = (a'_i, i = 1, \ldots, p), \ b' \in R, \ |b'| < \frac{1}{p-1} \sum_{(i,j) \in E} \sqrt{a'_i a'_j}\} . \tag{17}$$

We are now in a position to give the analytic expression of $I_{\mathcal{G}}(\delta, D)$.

**Theorem 4.1** *For $\mathcal{G} = T$ as described above, $\delta > 0$ and $D = m(a', b') \in P_{\mathcal{G}}^*$, the normalizing constant $I_{\mathcal{G}}(\delta, D)$ is finite and equal to*

$$I_{\mathcal{G}}(\delta, D) = 2^{\frac{\delta}{2} + p - 1} \Gamma(\frac{\delta}{2}) \left( \prod_{i=1}^{p} (a'_i)^{d_i - 2} \right)^{\frac{\delta}{4}} \int_{-\infty}^{\infty} \left( \prod_{(i,j) \in E} K_{\frac{\delta}{2}}(|b|\sqrt{a'_i a'_j}) \right) |b|^{\frac{p\delta}{2}} e^{-(p-1)bb'} db \tag{18}$$

12

where $d_i$ denotes the number of neighbours of the vertex $i$ in the tree $(V, E)$.

For $\delta = 1$, we have

$$I_{\mathcal{G}}(1, D) = (2\pi)^{\frac{p}{2}} \prod_{i=1}^{p} (a_i')^{-\frac{1}{2}} \Big( [\sum_{(i,j)\in E} (a_i' a_j') - (p-1)b']^{-1} + [\sum_{(i,j)\in E} (a_i' a_j') + (p-1)b']^{-1} \Big).$$

For $\delta = 3$, Let $\sigma_k$ the $k$th elementary function of the variables $\sqrt{a_i' a_j'}$, $(i, j) \in E$. We have

$$
\begin{aligned}
I_{\mathcal{G}}(3, D) \;=\;& 2^{\frac{p}{2}-1} \pi^{\frac{p}{2}} \prod_{i=1}^{p} (a_i')^{-\frac{3}{2}} \sum_{k=0}^{p-1} \sigma_k \Gamma(k+1) \Big( [\sum_{(i,j)\in E} (a_i' a_j')^{\frac{1}{2}} - (p-1)b']^{-(k+1)} \\
& -[\sum_{(i,j)\in E} (a_i' a_j')^{\frac{1}{2}} + (p-1)b']^{-(k+1)} \Big).
\end{aligned}
$$

## 4.3 The star graph with its $n$ leaves in one colour class

An example of star graph with its $n$ leaves in one color class and different colors for the edges and the central node is given in Figure 1(b). For $a \in R, c \in R, b = (b_1, \ldots, b_n) \in R^n$, let $L(\mathcal{G})$ the linear space of matrices of the form

$$
l(a, b, c) =
\begin{bmatrix}
a & b_1 & b_2 & \ldots & b_n \\
b_1 & c & 0 & \ldots & 0 \\
b_2 & 0 & c & \ldots & 0 \\
\ldots & \ldots & \ldots & \ldots & \ldots \\
b_n & 0 & 0 & \ldots & c
\end{bmatrix}.
$$

It is easy to see that the determinant of $l(a, b, c)$ is

$$|l(a, b, c)| = c^n \left( a - \frac{\|b\|^2}{c} \right) \tag{19}$$

and therefore, $P_{\mathcal{G}}$ is the open cone

$$P_{\mathcal{G}} = \{l(a, b, c) \in L(\mathcal{G}) : \ c > 0, \ a - \frac{\|b\|^2}{c} > 0.\}$$

The dual cone $P^*(\mathcal{G})$ and the normalizing constant $I_{\mathcal{G}}(\delta, D)$ are given below.

**Proposition 4.2** *For a star graph with all $n$ leaves in one colour class, the dual of $P_{\mathcal{G}}$ is*

$$P_{\mathcal{G}}^* = \{l(a', b', c') \in L(\mathcal{G}) \mid \|b'\|^2 \leq na'c'\}. \tag{20}$$

13

**Theorem 4.2** *For $\mathcal{G}$ a star graph with all $n$ leaves in the same colour class, $\delta \geq 1$ and $D = l(a', b', c') \in P_{\mathcal{G}}^*$, the normalizing constant of the CG-Wishart is*

$$I_{\mathcal{G}}(\delta, D) = 2^{\frac{\delta+n\delta+2}{2}} \pi^{n/2} \times a'^{(\frac{\delta}{2}-1)(n-1)} \times \frac{1}{(na'c' - \|b'\|^2)^{(\delta-1)\frac{n}{2}+1}} \times \Gamma((\delta-1)\frac{n}{2}+1)\Gamma(\frac{\delta}{2}).$$

### 4.4 The star graph with all vertices in one colour class

An example of star graph with all vertices in one color class and different colors for the edges is given in Figure 1(c). This case is a special case of the preceding one and therefore, we have immediately that

$$P_{\mathcal{G}} = \{l(a, b, a) \in L(\mathcal{G}) \mid a > 0, a^2 - \|b\|^2 > 0\}.$$

Since this is a well-known cone, called the Lorentz cone, we know also that it is self dual and therefore

$$P_{\mathcal{G}}^* = \{l(a', b', a') \in L(\mathcal{G}) \mid a' > 0, (a')^2 - \|b'\|^2 > 0\}.$$

It remains to compute $I_{\mathcal{G}}(\delta, D)$.

**Theorem 4.3** *For $\mathcal{G}$ the star graph with $n$ leaves and all vertices in the same colour class, $\delta > 0$ and $D = l(a', b', a') \in P_{\mathcal{G}}^*$, the normalizing constant of the CG-Wishart is*

$$I_{\mathcal{G}}(\delta, D) = \frac{2^{\frac{(n+1)\delta}{2}-1}C_n\Gamma((n+1)\frac{\delta}{2})}{(n+1)^{\frac{(n+1)\delta}{2}}(a')^{\frac{(n+1)\delta}{2}}}B(\frac{\delta}{2}, \frac{n}{2}) \, _2F_1\left((n+1)\frac{\delta}{4}, (n+1)\frac{\delta}{4}+\frac{1}{2}, \frac{n+\delta}{2}; u\right)$$

*where $u = \left(\frac{2\|b'\|}{(n+1)a'}\right)^2$ and $B(\frac{\delta}{2}, \frac{n}{2})$ is the Beta function with argument $(\frac{\delta}{2}, \frac{n}{2})$.*

### 4.5 A complete graph on three vertices with two edges in the same colour class

This graph is represented in Figure 1(d). In this case, the cone $P_{\mathcal{G}}$ is the set of positive definite matrices $K = (k_{ij})_{1 \leq i,j \leq 3}$ with $k_{13} = k_{23}$.

The dual cone $P^*(\mathcal{G})$ and the normalizing constant $I_{\mathcal{G}}(\delta, D)$ are given below.

**Proposition 4.3** *For the graph in Figure 1(d), the dual of $P_{\mathcal{G}}$ is*

$$P_{\mathcal{G}}^* = \{D = (d_{ij})_{1 \leq i,j \leq 3} \in S \mid d_{13} = d_{23},$$
$$d_{ii} > 0, i = 1, 2, 3, \ d_{12}^2 < d_{11}d_{22}, \ 4d_{13}^2 < (d_{11} + d_{22} + 2d_{12})d_{33}\}.$$
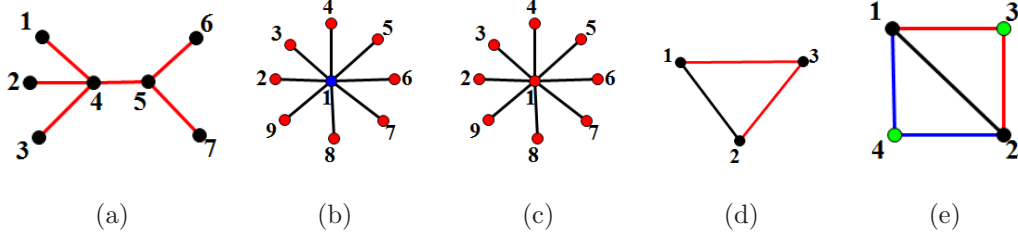
14

Figure 1: Black vertices and edges are all of different colours. (a) The colored tree. (b) The coloured star with the centre vertex of a different colour. (c) The coloured star with all vertices of the same colour. (d) The triangle with two edges of the same colour. (e) The decomposable graph with three different colours for the edges.

**Theorem 4.4** *For $\mathcal{G}$ as in Figure 1(d), $\delta > 0$ and $D \in P_{\mathcal{G}}^*$, the normalizing constant of the CG-Wishart is*

$$I_{\mathcal{G}}(\delta, D) = 2^{\frac{3\delta+4}{2}} \pi \Gamma(\frac{\delta}{2}) \left(\Gamma(\frac{\delta+1}{2})\right)^2 (d_{11} + d_{22} + 2d_{12})^{\frac{\delta}{2}} [d_{33}(d_{11} + d_{22} + 2d_{12}) - 4d_{13}^2]^{-\frac{\delta+1}{2}}$$
$$\times (d_{11}d_{22} - d_{12}^2)^{-\frac{\delta+1}{2}}.$$

## 4.6 A decomposable graph with three vertex classes and three edge classes

This graph is represented in Figure 1(e). Then the cone $P_{\mathcal{G}}$ is the set of matrices of the form

$$K = \begin{pmatrix} k_{11} & k_{12} & k_{13} & k_{14} \\ k_{12} & k_{22} & k_{13} & k_{14} \\ k_{13} & k_{13} & k_{33} & 0 \\ k_{14} & k_{14} & 0 & k_{33} \end{pmatrix}.$$

**Proposition 4.4** *For $\mathcal{G}$ as in Figure 1(e), the dual cone is the set of matrices*

$$P_{\mathcal{G}}^* = \{D = (d_{ij})_{1 \le i,j \le 4} \in S \mid d_{23} = d_{13}, d_{24} = d_{14}, d_{44} = d_{33}, \ d_{11} > 0,$$
$$d_{11}d_{22} - d_{12}^2 > 0, \ d_{11} + 2d_{12} + d_{22} > 0, \ d_{33}(d_{11} + 2d_{12} + d_{22}) - 2(d_{13}^2 + d_{14}^2) > 0\}.$$

**Theorem 4.5** *For $\mathcal{G}$ as in Figure 1(e), $\delta > 0$ and $D \in P_{\mathcal{G}}^*$, the normalizing constant of the CG-Wishart is*

$$
\begin{aligned}
I_{\mathcal{G}}(\delta, D) \;=\; & 2^{\delta+2} \pi^{\frac{3}{2}} \Gamma(\frac{\delta}{2}) \Gamma(\frac{\delta+1}{2}) \Gamma(\delta)(d_{11} + d_{22} + 2d_{12})^{\delta-1} (d_{11}d_{22} - d_{12}^2)^{-\frac{\delta+1}{2}} \\
& \times [d_{33}(d_{11} + d_{22} + 2d_{12}) - 2(d_{13}^2 + d_{14}^2)]^{-\delta}.
\end{aligned}
$$

# 5   Numerical experiments when we know the exact mean

In order to illustrate the performance of our MH algorithm, we conduct a numerical experiment for each of the colored graph (a) - (e) shown in Figure 1. In each case, for a given $D$ and $\delta$, we first derive $\log I_G(\delta, D)$, then the prior mean $E(K)$ under the CG-Wishart by differentiating $\log I_G(\delta, D)$ with respect to $-\frac{D}{2}$. We then sample from the CG-Wishart. We run the chain for 5000 iterations and discard the first 1000 samples as burn in. Our estimate $\hat{K}$ for $K$ is the average $\hat{K} = \frac{\sum_{i=1001}^{i=5000} \hat{K}_i}{4000}$ of the remaining 4000 iterations $\hat{K}_i, i = 1001, \ldots, 5000$. For arbitrary $K$ and $K'$ we define the normalized mean square error ($nmse$) between $K$ and $K'$ to be

$$
nmse(K, K') = \frac{||K - K'||_2^2}{||K'||_2^2}
$$

where $||K||_2^2$ is the sum of the squares of the entries of $K$. We repeat the previous experiment 100 times, obtain $\hat{K}^j, j = 1, \ldots, 100$ and compute

$$
\overline{nmse}(E(K), \hat{K}) = \frac{1}{100} \sum_{j=1}^{100} nmse(E(K), \hat{K}^j)
$$

where $E(K)$ is obtained by differentiation of $\log I_{\mathcal{G}}(\delta, D)$ with respect to $-\frac{D}{2}$ at our given $D$ and $\delta$.

For each graph in Figure 1, for an arbitrary $j \in \{1, \ldots, 100\}$, we give the trace plot of $\log |K_i^j|$, $i = 1000, \ldots, 5000$. The traceplot shows that the chain seems to be mixing well. We also provide the autocorrelation plot with time-lag $h$ for $\log |K_i^j|$, $i = 1000, \ldots, 5000$ in function of $h$ where, for an arbitrary given $j$, we define the autocorrelation coefficient for $Y_i = \log |K_i^j|$, $i = 1000, \ldots, 5000$ to be

$$
R_h = \frac{\sum_{i=1000}^{5000-h} (Y_i - \bar{Y})(Y_{i+h} - \bar{Y})}{\sum_{i=1000}^{5000} (Y_i - \bar{Y})^2}.
$$

| $\mathcal{G}$ | $\delta$ | $\log I_{\mathcal{G}}(\delta, D)$ | $\overline{nmse}(E(K), \hat{K})$ |
|---|---|---|---|
| Fig. 1(a) | 1 | $-\frac{1}{2}\sum_{i=1}^{7}\log a'_i + \log\left[\dfrac{1}{\sum_{i=1}^{6}(a'_i a'_{j_i})^{\frac{1}{2}}-6b'} - \dfrac{1}{\sum_{i=1}^{6}(a'_i a'_{j_i})^{\frac{1}{2}}+6b'}\right]$ | 0.0069 |
| Fig. 1(b) | 3 | $\frac{7}{2}\log a' - 9\log(8a'c' - \|b'\|^2)$ | 0.0187 |
| Fig. 1(c) | 3 | $-15\log a' + \log\ {}_2F_1\left(\frac{15}{2}, 8; 6; \frac{\|b'\|^2}{25a'^2}\right)$ | 0.0064 |
| Fig. 1(d) | 3 | $\frac{3}{2}\log d - 2\log(d_{33}d - 4d_{13}^2) - 2\log(d_{11}d_{22} - d_{12}^2)$ | 0.0005 |
| Fig. 1(e) | 3 | $2\log d - 3\log(d_{33}d - 2d_{13}^2 - 2d_{14}^2) - 2\log(d_{11}d_{22} - d_{12}^2)$ | 0.0009 |

Table 1: For the graphs of Fig. 1 and $\delta$ given: analytic expression of $\log I_{\mathcal{G}}(\delta, D)$ where $d = d_{11} + d_{22} + 2d_{12}$ and value of $\overline{nmse}(E(K), \hat{K})$ averaged over 100 experiments.

The autocorrelation plots indicate that the samples have a low autocorrelation. The numerical values of the matrices $D$, $E(K)$ and $\hat{K}$ as well as the traceplot and autocorrelation plot of $\log(|K|)$ for all five graphs in Figure 1 are given in Appendix 2 in the Supplementary file. An overview of calculations and results are given in Table 1 which, for all different five colored graphs in Figure 1, shows the parameter $\delta$ we chose for the prior distribution, $\log I_{\mathcal{G}}(\delta, D)$ and the normalized mean square errors. In order to obtain the mean $E(K)$ of the $CG$-Wishart for the graph in Figure 1(c), we use formula (2) to get the derivative of the hypergeometric function $_pF_q(a_1, \ldots, a_p; b_1, \ldots, b_q; z)$. We see that the normalized mean square error is of the order of $10^{-3}$ or less except for the star graph with all leaves of the same colour in Figure 1(b).
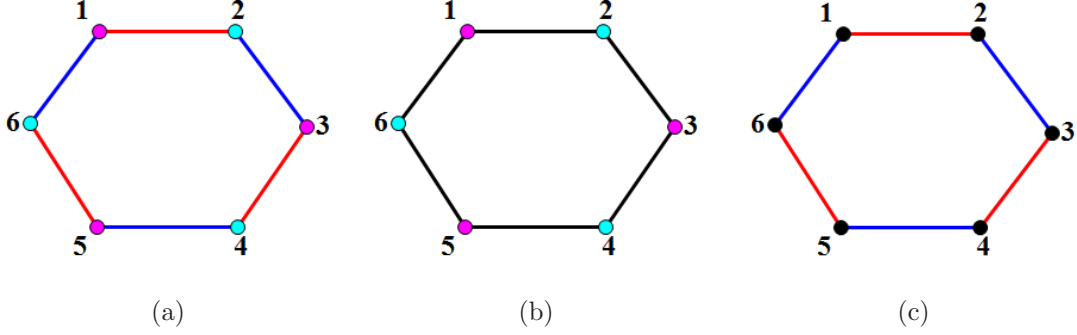
Figure 2: Cycles of length 6 with the three different patterns of colouring that we use for the cycles of length $p = 20$ and $p = 30$. Black vertices or edges indicate different arbitrary colours.

# 6  The posterior mean from simulated data: $p = 20$, $p = 30$

In this section, in order to assess the accuracy of our sampling method for larger graphs, we generate data from a $N(0, K^{-1})$ distribution with $K$ given in $P_{\mathcal{G}}$. We take the $CG$-Wishart with $\delta = 3$ and $D = I$ as the prior distribution of $K$. Clearly the posterior distribution will be $CG$-Wishart with parameters $\delta + n$ and $I + nS$ where $S$ is the sample covariance matrix. We will use this posterior and our sampling method of Section 3 to compute the posterior mean $E(K|S)$ as an estimate of $K$.

We run our experiment with six different coloured graphs. For three of them, the skeleton is a cycle of order $p = 20$ and for the other three, the skeleton is a cycle of order $p = 30$. For each cycle of order $p$, we give three different patterns of colouring which, for the sake of saving space, are illustrated in Figure 2 for $p = 6$. The values for the entries of $K$ for all three types of graphs are as follows:

$$K_{ii} = 0.1, \ i = 1, 3, \ldots, 2p - 1, \quad K_{ii} = 0.03. \ i = 2, 4, \ldots, p,$$
$$K_{i,i+1} = K_{i+1,i} = 0.01, \ i = 1, 2, \ldots, p - 1, \quad K_{1p} = K_{p1} = 0.01.$$

Though, for convenience, we chose, for all the models, the same values for the entries $K_{ij}, i \neq j$ to be all equal to .01 and for $K_{ii}$ to have two different values .1 and .03, in our computations, we used, of course, in each case, the model represented by each of the respective graphs.

For each graph, we generated 100 datasets from the $N(0, K^{-1})$ distribution. The posterior mean estimates are based on 5000 iterations after the first 1000 burn-in iterations. We denote $\hat{K} = (\hat{K}_{ij})_{1 \le i,j \le p}$ the posterior mean estimate.

Table 2 shows $\overline{nmse}(K, \hat{K})$ for the three colored models on the simulated examples when $p = 20$ and $p = 30$, averaged over 100 simulations. Standard errors are indicated in parentheses. Computations were performed on a 2 core 4 threads with i5-4200U, 2.3 GHZ chips and 8GB of RAM, running on Windows 8. We also give in Table 2 the average computing time per simulation in minutes.

Table 2: $\overline{nmse}(K, \hat{K})$ for the three colored models when $p = 20$ and $p = 30$.

| $\mathcal{G}$ | $p = 20$ | | $p = 30$ | |
| --- | --- | --- | --- | --- |
| | $\overline{nmse}(K, \hat{K})$ | Time/sim | $\overline{nmse}(K, \hat{K})$ | Time/sim |
| Fig.2 (a) | 0.005 (0.003) | 19.425 | 0.040 (0.021) | 86.423 |
| Fig.2 (b) | 0.011 (0.003) | 18.739 | 0.033 (0.011) | 82.876 |
| Fig.2 (c) | 0.039 (0.021) | 16.410 | 0.080 (0.033) | 82.563 |

In Table 3, for the graph of Fig. 2 (a) with $p = 20$ and $p = 30$, we give the values of the entries of $K$ together with their batch standard errors. For the other models, average

Table 3: The average estimates and batch standard errors for $K$ in Fig.2 (a).

| $p$ | $K_{11}$ | $K_{12}$ | $K_{1p}$ | $K_{22}$ |
| --- | --- | --- | --- | --- |
| 20 | 0.1040 (0.0005) | 0.0103 (0.0002) | 0.0104 (0.0002) | 0.0313 (0.0001) |
| 30 | 0.1223 (0.0009) | 0.0121 (0.0004) | 0.0125 (0.0004) | 0.0361 (0.0003) |

values of the entries together with batch standard errors are given in Appendix 3 in the Supplementary file.

*Remark 1.* At this point, we ought to make an important remark. In Section 5, we proved that the $CG$-Wishart was proper for $D \in P_{\mathcal{G}}^*$. When we compute the posterior mean in this section or more generally for any colored graph, even if $D > 0$ belongs to $P_{\mathcal{G}}^*$, the hyperparameter $D + nS$ does not usually belong to $P_{\mathcal{G}}^*$ of course and yet the integral

$I_{\mathcal{G}}(\delta + n, D + nS)$ converges. This is due to the fact that we can write $S$ as

$$S = S_1 + S_2$$

where $S_1$ is the projection of $S$ on the subspace of $p \times p$ matrices with fixed zeros according to $\mathcal{G}$ and equal entries for edges and vertices in the same colour class and $S_2$ belongs to its orthogonal complement. Since $K \in P_{\mathcal{G}}$, we have

$$0 < \langle K, S \rangle = \langle K, S_1 \rangle$$

and, since the inequality above is true for any $K \in P_{\mathcal{G}}$, it follows that $S_1$ belongs to $P_{\mathcal{G}}^*$. It follows also that $I_{\mathcal{G}}(\delta + n, D + nS)$ is finite.

*Remark 2.* For the computation of the posterior mean following our sampling scheme of Section 3, we may wonder whether we should take $Q$ to be such that $Q^t Q = (D + nS)^{-1}$ or $Q^t Q = (D + nS_1)^{-1}$. We take $Q$ to be such that $Q^t Q = (D + nS)^{-1}$ to use all the information given by the data.

# Appendix 1

## Proofs of Section 4

### Proof of Proposition 4.1
Let $M$ be the set of $p \times p$ matrices. Let

$$\mathcal{T}_G = \{X \in M \mid X_{ij} = 0, \text{ for } i < j, \ X_{ij} = s_{ij} \neq 0, \text{ for } i > j, (i,j) \in E, \ X_{ii} = t_i > 0, i = 1, \ldots, p\}$$

be the set of upper triangular matrices with positive diagonal elements and nonzero entries $X_{ij}, i > j$ only for $(i,j) \in E$. The vector $s = (s_{ij}, (i,j) \in E)$ belongs to $R^{p-1}$ since a tree with $p$ vertices has $p - 1$ edges and $t = (t_i, i = 1, \ldots, p)$ belongs to $R^p$. It is well-known (see Paulsen et al. [1989] and Roverato [2000]) that we can find a perfect elimination scheme enumeration of the vertices of $T$ such that, with this enumeration, $K \in P_{\mathcal{G}}$ can be written as $K = X(t,s)^T X(t,s)$ with $X(t,s) \in \mathcal{T}_G$. Then for $K = K(a,b)$ as in (16) we have

$$a_j = t_j^2 + \sum_{i \in E_j} s_{ij}^2, \ b = t_i s_{ij},$$

20

where $(t, s)$ is the Cholesky parametrization of $K \in P_{\mathcal{G}}$. We can also parametrize $K \in P_{\mathcal{G}}$ with $(t, b) \in (0, +\infty)^p \times R$ using

$$a_j = t_j^2 + b^2 \sum_{i \in E_j} \frac{1}{t_i^2}. \tag{21}$$

In this proof and the following one, we assume that the numbering of the vertices of $T$ follows a perfect elimination scheme ordering. We then say that the last vertex $p$ in that ordering is the root of the tree and we will write

$$E_j = \{i, i < j \mid (i, j) \in E\}.$$

For convenience, we denote by $C$ the right-hand side of equation (17).

We show first that $P_{\mathcal{G}}^* \subset C$. Let $D = m(a', b') \in P_{\mathcal{G}}^*$. Using (21), we have

$$\langle K, D \rangle = a_1 a_1' + \cdots + a_p a_p' + 2(p-1)bb' = t_p^2 a_p' + Ab^2 + 2Bb + C > 0, \tag{22}$$

where

$$A = \sum_{j=1}^{p} \left( \sum_{i \in E_j} \frac{1}{t_i^2} \right) a_j', \quad B = (p-1)b', \quad C = \sum_{i=1}^{p-1} t_i^2 a_i'. \tag{23}$$

Now observe that for fixed $i = 1, \ldots, p$ then either $i = p$ and the set $\{j; i \in E_j\}$ is empty since $p$ is the root of the tree, or the set $\{j; i \in E_j\}$ is reduced to one point, say $j_i$. Therefore we have $\sum_{i \in E_j} a_j' = a_{j_i}'$ for $i < p$ and zero for $i = p$. (For the graph in Figure 1 (a), we have $j_1 = j_2 = j_5 = j_6 = 7$ and $j_3 = j_4 = 6$) and it follows that

$$A = \sum_{i=1}^{p-1} \frac{1}{t_i^2} a_{j_i}'. \tag{24}$$

Let us prove that $a_j' > 0$ for all $j = 1, \ldots, p$. Take $(a_1, \ldots, a_p) \in [0, \infty)^p \setminus \{0, \ldots, 0\}$. Then $K(a, 0) \in \overline{P_{\mathcal{G}}} \setminus \{0\}$ and $\langle K, D \rangle = a_1 a_1' + \cdots + a_p a_p' > 0$ implies that $a_j' > 0$ for all $j$. Let us now prove that $Ab^2 + 2Bb + C \geq 0$ for all $b$. If not, there exists $b_0$ such that $Ab_0^2 + 2Bb_0 + C < 0$. Since $a_p' > 0$ taking $t_p$ very small and $b = b_0$ in (22) gives a contradiction.

Let us prove that

$$|b'| \leq \frac{1}{p-1} \sum_{(i,j) \in E} \sqrt{a_i' a_j'}. \tag{25}$$

21

Since $\forall b, \; Ab^2 + 2Bb + C \geq 0$, we have $B^2 \leq AC$. Now consider the function

$$(t_1, \ldots, t_{p-1}) \mapsto AC$$

and let us compute its minimum $A^*C^*$ on $(0, \infty)^p$. This function $AC$ is homogeneous of degree 0 and therefore if its minimum is reached at $t^* = (t_1^*, \ldots, t_{p-1}^*)$ it will also be reached on $\kappa t^*$ for any $\kappa > 0$. We have for $i = 1, \ldots, p - 1$

$$\frac{\partial}{\partial t_i} AC = 2 t_i a_i' A - \frac{2}{t_i^3} a_{j_i}' C = 0$$

and we therefore have

$$t_i^* = \kappa \left( \frac{a_{j_i}'}{a_i'} \right)^{1/4}, \quad A^* = C^* = \sum_{i=1}^{p-1} \sqrt{a_i' a_{j_i}'} = \sum_{(i,j) \in E} \sqrt{a_i' a_j'}.$$

Since $B^2 \leq AC$ for all $(t_1, \ldots, t_{p-1}) \in (0, \infty)^{p-1}$, we can claim that $B^2 \leq A^*C^*$ or equivalently (25).

Let us prove that inequality (25) is strict, that is $B^2 = A^*C^*$ is impossible. Suppose that $B^2 = A^*C^*$, i.e. $|b'| = A^*/(p-1) > 0$. Then with $t_i = t_i^*$ we get $Ab^2 + 2Bb + C = A^*(b + \operatorname{sign} b')^2$. Taking $b = -\operatorname{sign} b'$ and $t_i = t_i^*, i = 1, \ldots, p-1$ yields $Ab^2 + 2Bb + C = 0$. Now, letting also $t_p = 0$ in (22), we see that the left hand side of (22) is zero for an $(a, b) \in \overline{S} \setminus \{0\}$ which is not zero, since $b = \pm 1$. But this cannot happen for $D(a', b') \in P_{\mathcal{G}}^*$. Therefore (25) is strict and the proof of $P_{\mathcal{G}}^* \subset C$ is completed.

Let us now show that $C \subset P_{\mathcal{G}}^*$. For $D(a', b') \in C$ given, we want to show that $\langle K, D \rangle$ is positive for all $K(a, b) \in \overline{P_{\mathcal{G}}} \setminus \{0\}$. We will do so first for $K(a, b) \in P_{\mathcal{G}}$ and then for $K(a, b) \in \overline{P_{\mathcal{G}}} \setminus (P_{\mathcal{G}} \cup \{0\})$. For $K(a, b) \in P_{\mathcal{G}}$, $t_k > 0$ and $b \in R$. From (22), we have

$$\langle K, D \rangle = t_p^2 a_p' + Ab^2 + 2Bb + C = t_p^2 a_p' + A \left[ \left( b + \frac{B}{A} \right)^2 + \frac{1}{A^2} (AC - B^2) \right].$$

We have checked above that $AC - B^2 \geq 0$. Moreover $a_p' > 0$ since $D(a', b') \in C$. It follows immediately that $\langle K, D \rangle > 0$.

Let us now show $\langle K, D \rangle > 0$ for $K(a, b) \in \overline{P_{\mathcal{G}}} \setminus (P_{\mathcal{G}} \cup \{0\})$ that is for $t_1 \ldots t_p = 0$ and $(t_1, \ldots, t_p, b) \neq 0$. We need only show that then $\langle K, D \rangle \neq 0$. But $0 = a_1 a_1' + \cdots + a_p a_p' + 2(p-1)bb' = \sum_{i=1}^p t_i^2 a_i'$ implies that $t_i = 0$ for all $i = 1, \ldots, p$ since $(a', b') \in C$ implies $a_i' > 0$. But

22

since $b = t_i s_{ij}$, this implies $b = 0$ but this is impossible since we exclude the zero matrix for $K$.

**Proof of Theorem 4.1**

In $I_{\mathcal{G}}(\delta, D)$ we make the change of variable (21). Switching to these Cholesky coordinates leads to the Jacobian $dadb = 2^p t_1 \ldots t_p \, dbdt$. As seen before the new domain of integration is the product

$$\{(b, t); t_k > 0, b \in R\} = (0, \infty)^p \times R.$$

With the notation $A, B, C$ of (23), we have

$$\langle K(a, b), D(a', b') \rangle = 2(p-1)bb' + a_1 a_1' + \cdots + a_p a_p' = t_p^2 a_p' + Ab^2 + 2Bb + C.$$

Using (24) for the expression of $A$, we obtain

$$
\begin{aligned}
I_G(\delta, D) &= 2^p \int_{(0,\infty)^p \times R} (t_1 \ldots t_p)^{\delta-1} e^{-(p-1)bb'} e^{-\frac{t_p^2 a_p'}{2}} \prod_{i=1}^{p-1} e^{-\frac{t_i^2 a_i'}{2} - \frac{b^2 a_{j_i}'}{2t_i^2}} dt_1 \ldots dt_p \, db \\
&= 2^p \int_0^\infty e^{-\frac{t_p^2 a_p'}{2}} t_p^{\delta-1} dt_p \int_{-\infty}^\infty e^{-(p-1)bb'} \prod_{i=1}^{p-1} \left( K_{\delta/2}(|b|(a_i' a_{j_i}')^{1/2})(|b|\sqrt{a_{j_i}'/a_i})^{\delta/2} \right) db \\
&= 2^{p+\frac{\delta}{2}-1} \frac{\Gamma(\delta/2)}{(a_p')^{\delta/2}} \left( \prod_{i=1}^{p-1} \frac{a_{j_i}'}{a_i'} \right)^{\delta/4} J_\delta(D) \quad (26)
\end{aligned}
$$

with the notation

$$J_\delta(D) = \int_{-\infty}^\infty e^{-(p-1)bb'} |b|^{(p-1)\delta/2} \prod_{i=1}^{p-1} K_{\delta/2}(|b|(a_i' a_{j_i}')^{1/2}) db. \quad (27)$$

We now prove by induction that

$$\frac{1}{(a_p')^2} \times \prod_{i=1}^{p-1} \frac{a_{j_i}'}{a_i} = \prod_{i=1}^p (a_i')^{d_i - 2}. \quad (28)$$

Of course (28) is correct for $p = 2$. Suppose that (28) is true for any rooted tree with size $p$. Consider a rooted tree $T^*$ with vertices $\{0, 1, \ldots, p\}$ and root $p$ and numbered, as usual, such that $i \prec j$ implies $i \leq j$. Denote $T$ the induced tree with vertices $\{1, \ldots, p\}$. Finally

23

denote $d^* = (d_0^*, \ldots, d_p^*)$ and $d = (d_1, \ldots, d_p)$ the number of neighbours in $T^*$ and $T$. Then $d_0^* = 1$, $d_{j_0}^* = 1 + d_{j_0}$ and $d_i^* = d_i$ if $i \neq 0$ and $i \neq j_0$. This implies that

$$\frac{1}{(a_p')^2} \times \prod_{i=0}^{p-1} \frac{a_{j_i}'}{a_i} = \frac{a_{j_0}'}{a_0} \frac{1}{(a_p')^2} \times \prod_{i=1}^{p-1} \frac{a_{j_i}'}{a_i} \overset{(1)}{=} \frac{a_{j_0}'}{a_0} \prod_{i=1}^{p} (a_i')^{d_i - 2} \overset{(2)}{=} \prod_{i=0}^{p} (a_i')^{d_i^* - 2},$$

where (1) comes from the induction hypothesis and (2) from the link between $d$ and $d^*$. The induction hypothesis is extended to $p + 1$ and (28) is proved.

We now prove that $J_\delta(D)$ defined by (27) converges if $D = m(a', b') \in P_{\mathcal{G}}^*$ where $P_{\mathcal{G}}^*$ is the convex cone defined in Proposition 3. We write $J_\delta(D)$ as the sum

$$J_\delta(D) = \int_{-\infty}^{0} \ldots db + \int_{0}^{+\infty} \ldots db. \tag{29}$$

When $b \to \pm\infty$, $|b| \to +\infty$. From Watson [1995] page 202, 7.23 (1) we have

$$K_\lambda(s) \sim_{s \to \infty} \sqrt{\frac{\pi}{2}} \frac{e^{-s}}{s^{1/2}}.$$

We use this fact to analyse the convergence of $J_\delta(D)$. If $D = m(a', b') \in P_{\mathcal{G}}^*$, from the asymptotic formula above, we see that the integrands in both integral on the RHS of (29), when $|b|$ goes to infinity, behave like $|b|^c e^{-|b|H}$ where, since $m(a', b') \in P_{\mathcal{G}}^*$,

$$H = \sum_{(i,j) \in E}^{p} \sqrt{a_i' a_j'} - (p-1)|b'|\text{sign}(bb') > 0$$

and $c = (p-1)\frac{\delta-1}{2}$. Since the argument of (27) is continuous, both integrals converge at infinity.

To study the convergence of these integrals when $b \to 0$, we recall that

$$2K_\lambda(s) = \int_{0}^{+\infty} x^{\lambda-1} e^{-\frac{s}{2}(x+\frac{1}{x})} dx.$$

Making the change of variable $u = sx$ in the expression of $2K_\lambda(s)$ we see that

$$K_\lambda(s) \sim_{s \to 0} s^{-\lambda} 2^{\lambda-1} \Gamma(\lambda).$$

Therefore, for both integrals in the RHS of (29), the integrand is equivalent to

$$\left(|b|^{-\frac{\delta}{2}} \Gamma(\frac{\delta}{2})\right)^{p-1} |b|^{\frac{p\delta}{2}} e^{-(p-1)bb'} = |b|^{\frac{\delta}{2}} e^{-(p-1)bb'}$$

24

and therefore both integrals converge at 0. The expression (18) of the normalizing constant is now proved.

By (18), $I_G(1, D) = 2^{p-\frac{1}{2}} \Gamma(\frac{1}{2})(a'_p)^{-\frac{1}{2}} (\prod_{i=1}^{p-1} \frac{a'_{j_i}}{a'_i})^{\frac{1}{4}} J_1(D)$, where

$$
\begin{aligned}
J_1(D) &= \int_{-\infty}^{\infty} e^{-(p-1)bb'} |b|^{\frac{p-1}{2}} \prod_{i=1}^{p-1} K_{\frac{1}{2}}(|b|(a'_i a'_{j_i})^{\frac{1}{2}}) db \\
&= \int_{-\infty}^{\infty} e^{-(p-1)bb'} |b|^{\frac{p-1}{2}} (\prod_{i=1}^{p-1} \sqrt{\frac{\pi}{2}} |b|^{-\frac{1}{2}} (a'_i a'_{j_i})^{-\frac{1}{4}} e^{|b|(a'_i a'_{j_i})^{\frac{1}{2}}}) db \\
&= (\frac{\pi}{2})^{\frac{p-1}{2}} \prod_{i=1}^{p-1} (a'_i a'_{j_i})^{-\frac{1}{4}} \int_{-\infty}^{\infty} e^{-(p-1)bb' - |b| \sum_{i=1}^{p-1} (a'_i a'_{j_i})^{\frac{1}{2}}} db.
\end{aligned}
$$

We compute the integral

$$
\begin{aligned}
\int_{-\infty}^{\infty} e^{-(p-1)bb' - |b| \sum_{i=1}^{p-1} (a'_i a'_{j_i})^{\frac{1}{2}}} db &= \int_{-\infty}^{0} e^{-(p-1)bb' + b \sum_{i=1}^{p-1} (a'_i a'_{j_i})^{\frac{1}{2}}} db + \int_{0}^{\infty} e^{-(p-1)bb' - b \sum_{i=1}^{p-1} (a'_i a'_{j_i})^{\frac{1}{2}}} db \\
&= \frac{1}{\sum_{i=1}^{p-1} (a'_i a'_{j_i}) - (p-1)b'} + \frac{1}{\sum_{i=1}^{p-1} (a'_i a'_{j_i}) + (p-1)b'}.
\end{aligned}
$$

Therefore

$$
J_1(D) = (\frac{\pi}{2})^{\frac{p-1}{2}} \prod_{i=1}^{p-1} (a'_i a'_{j_i})^{-\frac{1}{4}} \left[ \left( \sum_{i=1}^{p-1} (a'_i a'_{j_i})^{\frac{1}{2}} - (p-1)b' \right)^{-1} + \left( \sum_{i=1}^{p-1} (a'_i a'_{j_i})^{\frac{1}{2}} + (p-1)b' \right)^{-1} \right].
$$

Since $\sum_{i=1}^{p-1} (a'_i a'_{j_i}) = \sum_{(i,j) \in E} (a'_i a'_j)$, this yields the expression of $I_G(1, D)$.

Similarly, from (18), $I_G(3, D) = 2^{p+\frac{1}{2}}\Gamma(\frac{3}{2})(a'_p)^{-\frac{3}{2}}\prod_{i=1}^{p-1}(\frac{a'_{j_i}}{a'_i})^{-\frac{3}{4}}J_3(D)$ with

$$
\begin{aligned}
J_3(D) &= \int_{-\infty}^{\infty} e^{-(p-1)bb'}|b|^{\frac{3}{2}(p-1)}\prod_{i=1}^{p-1}K_{\frac{3}{2}}(|b|(a'_i a'_{j_i})^{\frac{1}{2}})db \\
&= \int_{-\infty}^{\infty} e^{-(p-1)bb'}|b|^{\frac{3}{2}(p-1)}\prod_{i=1}^{p-1}[\sqrt{\frac{\pi}{2}}(|b|^{-\frac{1}{2}}(a'_i a'_{j_i})^{-\frac{1}{4}} + |b|^{-\frac{3}{2}}(a'_i a'_{j_i})^{-\frac{3}{4}})e^{-|b|(a'_i a'_{j_i})^{\frac{1}{2}}}]db \\
&= \int_{-\infty}^{\infty} e^{-(p-1)bb'}|b|^{\frac{3}{2}(p-1)}(\frac{\pi}{2})^{\frac{p-1}{2}}|b|^{-\frac{3}{2}(p-1)}\prod_{i=1}^{p-1}(a'_i a'_{j_i})^{-\frac{3}{4}}\prod_{i=1}^{p-1}[(|b|(a'_i a'_{j_i})^{\frac{1}{2}} + 1)e^{-|b|(a'_i a'_{j_i})^{\frac{1}{2}}}]db \\
&= (\frac{\pi}{2})^{\frac{p-1}{2}}\prod_{i=1}^{p-1}(a'_i a'_{j_i})^{-\frac{3}{4}}\int_{-\infty}^{\infty} e^{-(p-1)bb'-|b|\sum_{i=1}^{p-1}(a'_i a'_{j_i})^{\frac{1}{2}}}\prod_{i=1}^{p-1}(1+|b|(a'_i a'_{j_i})^{\frac{1}{2}})db \\
&= (\frac{\pi}{2})^{\frac{p-1}{2}}\prod_{i=1}^{p-1}(a'_i a'_{j_i})^{-\frac{3}{4}}\int_{-\infty}^{\infty} e^{-(p-1)bb'-|b|\sum_{i=1}^{p-1}(a'_i a'_{j_i})^{\frac{1}{2}}}(1+|b|\sigma_1+|b|^2\sigma_2+\ldots+|b|^{p-1}\sigma_{p-1})db \\
&= (\frac{\pi}{2})^{\frac{p-1}{2}}\prod_{i=1}^{p-1}(a'_i a'_{j_i})^{-\frac{3}{4}}\sum_{k=0}^{p-1}\sigma_k\int_{-\infty}^{\infty} e^{-(p-1)bb'-|b|\sum_{i=1}^{p-1}(a'_i a'_{j_i})^{\frac{1}{2}}}|b|^k db\,,
\end{aligned}
$$

where the $\sigma_i = \sigma_i(\sqrt{a'_i a'_{j_i}}, i = 1,\ldots,p-1)$ are the symmetric functions of $\sqrt{a'_i a'_{j_i}}, i = 1,\ldots,p-1$. Since

$$
\int_{-\infty}^{\infty} e^{-(p-1)bb'-|b|\sum_{i=1}^{p-1}(a'_i a'_{j_i})^{\frac{1}{2}}}|b|^{m-1}db = \Gamma(m)\left[\left(\sum_{i=1}^{p-1}(a'_i a'_{j_i})^{\frac{1}{2}} - (p-1)b'\right)^{-m} - \left(\sum_{i=1}^{p-1}(a'_i a'_{j_i})^{\frac{1}{2}} + (p-1)b'\right)^{-m}\right],
$$

then

$$
I_G(3,D) = 2^{\frac{p}{2}-1}\pi^{\frac{p}{2}}\prod_{i=1}^{p}(a'_i)^{-\frac{3}{2}}\sum_{k=0}^{p-1}\sigma_k\Gamma(k+1)\left[\left(\sum_{i=1}^{p-1}(a'_i a'_{j_i})^{\frac{1}{2}} - (p-1)b'\right)^{-k-1} - \left(\sum_{i=1}^{p-1}(a'_i a'_{j_i})^{\frac{1}{2}} + (p-1)b'\right)^{-k-1}\right].
$$

This yields the expression of $I_G(3, D)$.

**Proof of Proposition 4.2**

By definition $P_G^* = \{D = l(a', b', c') \in M(\mathcal{G}) \mid \langle K, D\rangle > 0, \; K \in \bar{P}_{\mathcal{G}} \setminus \{0\}\}$. Let $\beta$ denote the angle between $b$ and $b'$. Then, since $\cos\beta > -1$

$$
\langle K, D\rangle = aa' + ncc' + 2||b||||b'||\cos\beta > aa' + ncc' - 2||b||||b'||.
$$

Therefore $2||b||||b'|| < aa' + ncc'$ and since $ac > 0$, $\frac{4||b||^2||b'||^2}{ac} < \frac{(aa'+ncc')^2}{ac}$. By differentiation with respect to $a$ and $c$, we see that $\frac{(aa'+ncc')^2}{ac} \geq 4na'c'$ and therefore $\langle K, D\rangle > 0$ implies that $||b'||^2 < na'c'$.

**Proof of Theorem 4.2.** Let us introduce the matrix

$$A(r,s,t) = \begin{bmatrix} r & s_1 & s_2 & \dots & s_n \\ 0 & t & 0 & \dots & 0 \\ 0 & 0 & t & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & t \end{bmatrix}.$$

If $(a,b,c) \in P_{\mathcal{G}}$ the only triple $(r,s,t)$ such that $t > 0$ and $r > 0$ and such that

$$K(a,b,c) = A(r,s,t)A^T(r,s,t) = \begin{bmatrix} r^2 + \|s\|^2 & s't \\ ts & t^2 I_n \end{bmatrix}$$

satisfies $r = (a - \frac{\|b\|^2}{c})^{1/2}$, $t = \sqrt{c}$, $s = \frac{b}{\sqrt{c}}$. A new parameterization of $P_{\mathcal{G}}$ is therefore given by the change of variables $(a,b,c)$ into $(r,s,t)$ with $a = r^2 + \|s\|^2, b = ts, \ c = t^2$, where $(r,s,t)$ belongs to

$$\{(r,s,t); r > 0, s \in R^n, t > 0\} = (0,\infty) \times R^n \times (0,\infty).$$

With this parameterization, from (19), we have $\det K = r^2 t^{2n}$ and $dadbdc = 4rt^{n+1}drdsdt$. Then

$$
\begin{aligned}
I_{\mathcal{G}}(\delta, D) &= 4 \int_0^\infty \int_0^\infty \left( \int_{R^n} e^{\frac{-\|s\|^2 a' - 2t\langle s, b'\rangle}{2}} ds \right) r^{\delta-1} t^{(\delta-1)n+1} e^{\frac{-r^2 a' - nt^2 c'}{2}} drdt \\
&= 4 \left(\frac{\pi}{a'}\right)^{n/2} \int_0^\infty e^{\frac{-nt^2 c'}{2} + \frac{t^2 \|b'\|^2}{2a'}} t^{(\delta-1)n+1} dt \times \int_0^\infty r^{\delta-1} e^{\frac{-r^2 a'}{2}} dr \\
&= \left(\frac{\pi}{a'}\right)^{n/2} \int_0^\infty e^{\frac{-nvc'}{2} + \frac{v\|b'\|^2}{2a'}} v^{(\delta-1)\frac{n}{2}} dv \times \int_0^\infty v^{\frac{\delta}{2}-1} e^{\frac{-va'}{2}} dv \\
&= 2^{\frac{\delta+n\delta+2}{2}} \pi^{n/2} a'^{(\frac{\delta}{2}-1)(n-1)} \frac{1}{(na'c' - \|b'\|^2)^{(\delta-1)\frac{n}{2}+1}} \Gamma((\delta-1)\frac{n}{2}+1)\Gamma(\frac{\delta}{2}).
\end{aligned}
$$

**Proof of Theorem 4.3.**

$$
\begin{aligned}
I_{\mathcal{G}}(\delta, D) &= \int_{R^n} \left( \int_{\|b\|}^\infty a^{n\frac{\delta-2}{2}} (a - \frac{\|b\|^2}{a})^{\frac{\delta-2}{2}} \exp -\frac{1}{2}\{(n+1)aa' + 2\langle b, b'\rangle\} da \right) db \\
&= \int_{R^n} \left( \int_{\|b\|}^\infty a^{(n-1)\frac{\delta-2}{2}} (a^2 - \|b\|^2)^{\frac{\delta-2}{2}} \exp -\frac{1}{2}\{(n+1)aa' + 2\langle b, b'\rangle\} da \right) db
\end{aligned}
$$

27

Let us make the change of variable

$$(a, b) \in (||b||, +\infty) \times R^n \mapsto (u, R, \theta) \in (0, 1) \times (0, +\infty) \times S$$

where $b = R\theta$ and $S$ is the unit sphere in $R^n$ and $a = \frac{R}{\sqrt{u}}$. We have $da\,db = -\frac{1}{2u^{3/2}}RC_n R^{n-1}da\,dR\,d\theta$ where $C_n$ is the surface area of $S$. Then

$$
\begin{aligned}
I_{\mathcal{G}}(\delta, D) &= \frac{C_n}{2}\int_S \left[\int_0^{+\infty}\left(\int_0^1 R^{(n-1)\frac{\delta-2}{2}}u^{-(n-1)\frac{\delta-2}{4}}R^{\delta-2}(\frac{1}{u}-1)^{\frac{\delta-2}{2}}\exp-\{\frac{(n+1)Ra'}{2\sqrt{u}}+R\langle\theta,b'\rangle\}u^{-3/2}du\right)R^n\,dR\right]d\theta \\
&= \frac{C_n}{2}\int_S\left[\int_0^{+\infty}\left(\int_0^1 R^{(n-1)\frac{\delta-2}{2}}u^{-(n+1)\frac{\delta-2}{4}}R^{\delta-2}(1-u)^{\frac{\delta-2}{2}}\exp-\{\frac{(n+1)Ra'}{2\sqrt{u}}+R\langle\theta,b'\rangle\}u^{-3/2}du\right)R^n\,dR\right]d\theta \\
&= \frac{C_n}{2}\int_S\left[\int_0^{+\infty}\left(\int_0^1 R^{(n+1)\frac{\delta}{2}-1}u^{-(n+1)\frac{\delta-2}{4}-\frac{3}{2}}(1-u)^{\frac{\delta-2}{2}}\exp-R\{\frac{(n+1)a'}{2\sqrt{u}}+\langle\theta,b'\rangle\}du\right)dR\right]d\theta \\
&= \frac{C_n}{2}\int_S\left[\int_0^1 u^{-(n+1)\frac{\delta-2}{4}-\frac{3}{2}}(1-u)^{\frac{\delta-2}{2}}\left(\int_0^{+\infty}R^{(n+1)\frac{\delta}{2}-1}\exp-R\{\frac{(n+1)a'}{2\sqrt{u}}+\langle\theta,b'\rangle\}dR\right)du\right]d\theta \\
&= \frac{C_n\Gamma((n+1)\frac{\delta}{2})}{2}\int_S\left[\int_0^1 u^{-(n+1)\frac{\delta-2}{4}-\frac{3}{2}}(1-u)^{\frac{\delta-2}{2}}\left(\frac{(n+1)a'}{2\sqrt{u}}+\langle\theta,b'\rangle\right)^{-(n+1)\frac{\delta}{2}}du\right]d\theta \\
&= \frac{C_n\Gamma((n+1)\frac{\delta}{2})}{2}(\frac{(n+1)a'}{2})^{-(n+1)\frac{\delta}{2}}\int_S\left[\int_0^1 u^{-(n+1)\frac{\delta-2}{4}-\frac{3}{2}}(1-u)^{\frac{\delta-2}{2}}u^{(n+1)\frac{\delta}{4}}\left(1+\frac{2}{(n+1)a'}\sqrt{u}\langle\theta,b'\rangle\right)^{-(n+1)\frac{\delta}{2}}du\right]d\theta \\
&= K_{n,\delta}(a')\int_S\left[\int_0^1 u^{\frac{n}{2}-1}(1-u)^{\frac{\delta}{2}-1}\sum_{k=0}^{\infty}(-1)^k\left(\frac{2\langle\theta,b'\rangle}{(n+1)a'}\right)^k u^{\frac{k}{2}}\frac{\left((n+1)\frac{\delta}{2}\right)_k}{k!}\,du\right]d\theta
\end{aligned}
$$

where $K_{n,\delta}(a') = \dfrac{2^{\frac{(n+1)\delta}{2}-1}C_n\Gamma((n+1)\frac{\delta}{2})}{(n+1)^{\frac{(n+1)\delta}{2}}(a')^{\frac{(n+1)\delta}{2}}}$. Therefore

$$
\begin{aligned}
I_{\mathcal{G}}(\delta, D) &= K_{n,\delta}(a')\sum_{k=0}^{\infty}(-1)^k\left(\frac{2}{(n+1)a'}\right)^k\frac{\left((n+1)\frac{\delta}{2}\right)_k}{k!}\int_0^1 u^{\frac{k+n}{2}-1}(1-u)^{\frac{\delta}{2}-1}du\int_S\langle\theta,b'\rangle^k d\theta \\
&= K_{n,\delta}(a')\sum_{k=0}^{\infty}\left(\frac{2}{(n+1)a'}\right)^{2k}\frac{\left((n+1)\frac{\delta}{2}\right)_{2k}}{(2k)!}\int_0^1 u^{\frac{2k+n}{2}-1}(1-u)^{\frac{\delta}{2}-1}du\int_S\langle\theta,b'\rangle^{2k}d\theta \\
&= K_{n,\delta}(a')\sum_{k=0}^{\infty}\left(\frac{2}{(n+1)a'}\right)^{2k}\frac{\left((n+1)\frac{\delta}{2}\right)_{2k}}{(2k)!}\frac{\Gamma(k+\frac{n}{2})\Gamma(\frac{\delta}{2})}{\Gamma(k+\frac{\delta+n}{2})}||b'||^{2k}\frac{(1/2)_k}{(n/2)_k}
\end{aligned}
$$

We now use the fact that $(\alpha)_{2k} = 2^{2k}\left(\frac{\alpha}{2}\right)_k\left(\frac{\alpha+1}{2}\right)_k$ and $\Gamma(\alpha+k) = \Gamma(\alpha)(\alpha)_k$. We also use the fact that

$$(2k)! = (135...(2k-1))(246...2k) = 2^k k!2^k\frac{1}{2}\frac{3}{2}...\frac{2k-1}{2} = 2^{2k}k!\frac{1}{2}(\frac{1}{2}+1)(\frac{1}{2}+2)...(\frac{1}{2}+(k-1)) = 2^{2k}k!\left(\frac{1}{2}\right)_k.$$

Finally, since the integral is rotational symmetric, we take $b' = ||b'||e_1$ so that $\langle\theta,b'\rangle = \theta_1||b'||$ and recalling that $d\theta$ is the distribution of $\frac{Z}{||Z||}$ when $Z \sim N(0,1)$ so that $\theta_1 = \frac{Z_1}{\sqrt{Z_1^2+...+Z_n^2}}$ which is then such that $\theta_1^2 \sim \text{Beta}(\frac{1}{2}, \frac{n-1}{2})$, for $v = \theta_1$, we have

$$\int_S\langle\theta,b'\rangle^{2k}d\theta = \frac{||b'||^{2k}}{B(\frac{1}{2}, \frac{n-1}{2})}\int_0^1 v^{k-\frac{1}{2}}(1-v)^{\frac{n-1}{2}-1}dv = ||b'||^{2k}\frac{(1/2)_k}{(n/2)_k}.$$

Writing $B(\alpha, \beta)$ for the Beta function with argument $(\alpha, \beta)$, we obtain

$$
\begin{aligned}
I_{\mathcal{G}}(\delta, D) &= K_{n,\delta}(a')B(\frac{\delta}{2}, \frac{n}{2}) \sum_{k=0}^{\infty} \left(\frac{2}{(n+1)a'}\right)^{2k} \frac{2^{2k}}{(2k)!} \left((n+1)\frac{\delta}{4}\right)_k \left((n+1)\frac{\delta}{4} + \frac{1}{2}\right)_k \frac{\left(\frac{n}{2}\right)_k}{\left(\frac{n+\delta}{2}\right)_k} \|b'\|^{2k} \frac{(1/2)_k}{(n/2)_k} \\
&= K_{n,\delta}(a')B(\frac{\delta}{2}, \frac{n}{2}) \sum_{k=0}^{\infty} \left(\frac{2}{(n+1)a'}\right)^{2k} \frac{2^{2k}}{2^{2k}k!(\frac{1}{2})_k} \left((n+1)\frac{\delta}{4}\right)_k \left((n+1)\frac{\delta}{4} + \frac{1}{2}\right)_k \frac{\left(\frac{n}{2}\right)_k}{\left(\frac{n+\delta}{2}\right)_k} \|b'\|^{2k} \frac{(1/2)_k}{(n/2)_k} .
\end{aligned}
$$

Let $u = \left(\frac{2\|b'\|}{(n+1)a'}\right)^2$. We note that since $D = l(a', b', a') \in P_{\mathcal{G}}^*$, then $u \leq 1$. After obvious simplifications in the expression above, we have

$$
\begin{aligned}
I_{\mathcal{G}}(\delta, D) &= K_{n,\delta}(a')B(\frac{\delta}{2}, \frac{n}{2}) \sum_{k=0}^{\infty} \frac{u^k}{k!} \frac{\left((n+1)\frac{\delta}{4}\right)_k \left((n+1)\frac{\delta}{4} + \frac{1}{2}\right)_k}{\left(\frac{n+\delta}{2}\right)_k} \\
&= K_{n,\delta}(a')B(\frac{\delta}{2}, \frac{n}{2}) \, {}_2F_1\left((n+1)\frac{\delta}{4}, (n+1)\frac{\delta}{4} + \frac{1}{2}, \frac{n+\delta}{2}; u\right) .
\end{aligned}
$$

**Proof of Proposition 4.3**

We write the Cholesky decomposition of $K$ under the form $K = AA^t$ with

$$
A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix} .
$$

Expressing the $k_{ij}$ in terms of the $a_{ij}$ and imposing $k_{13} = k_{23}$ immediately shows that we must have $a_{13} = a_{23}$. Then, let $D = (d_{ij})_{1 \leq i,j \leq 3}$ with $d_{13} = d_{23}$ since the dual of $P_{\mathcal{G}}$ must be in the same linear space as $P_{\mathcal{G}}$.

$$
\begin{aligned}
\langle K, D \rangle &= (a_{11}^1 + a_{12}^2 + a_{13}^2)d_{11} + (a_{22}^2 + a_{13}^2)d_{22} + a_{33}^2 d_{33} + 2(a_{22}a_{12} + a_{13}^2)d_{12} + 4a_{13}a_{33}d_{13} \\
&= a_{13}^2(d_{11} + d_{22} + 2d_{12}) + 4a_{13}a_{33}d_{13} + a_{12}^2 d_{11} + 2a_{12}a_{22}d_{12} + a_{11}^2 d_{11} + a_{22}^2 d_{22} + a_{33}^2 d_{33} ,
\end{aligned}
$$

which we view as a quadratic form $a^t M a$ with $a^t = (a_{13}, a_{33}, a_{12}, a_{22}, a_{33})$ and

$$
M = \begin{pmatrix} d_{11} + d_{22} + 2d_{12} & 2d_{13} & 0 & 0 & 0 \\ 2d_{13} & d_{33} & 0 & 0 & 0 \\ 0 & 0 & d_{11} & d_{12} & 0 \\ 0 & 0 & d_{12} & d_{22} & 0 \\ 0 & 0 & 0 & 0 & d_{11} \end{pmatrix} .
$$

29

Since $AA^t$ is the Cholesky parametrization of $P_{\mathcal{G}}$, clearly $K \in P_{\mathcal{G}}$ if and only if $a_{ii} > 0, i = 1, 2, 3$. If we can prove the following lemma, the condition $M > 0$ will yield the dual cone $P_{\mathcal{G}}^*$.

*Lemma A1.* The trace $\langle K, D \rangle$ is positive for all $K \in \bar{P}_{\mathcal{G}} \setminus \{0\}$ if and only if the matrix $M$ of the quadratic form $\langle K, D \rangle = a^t M a$ is positive definite

Let us now prove the lemma. Clearly if $M > 0$ then $\langle K, D \rangle = a^t M a > 0$ for all $a \in R^5$ and in particular for all $a$ with $a_{ii} > 0, i = 1, 2, 3$. Conversely let $a \in R^5$. Then $a$ can be written as

$$a = (\epsilon_1 a_{11}, \epsilon_2 a_{22}, \epsilon_3 a_{33}, a_{12}, a_{13})^t$$

where $\epsilon_i$ is the sign of $a_{ii}, i = 1, 2, 3$ and we have

$$a^t M a = (a_{11}^1 + a_{12}^2 + a_{13}^2) d_{11} + (a_{22}^2 + a_{13}^2) d_{22} + a_{33}^2 d_{33} + 2(\epsilon_2 a_{22} a_{12} + a_{13}^2) d_{12} + 4\epsilon_3 a_{13} a_{33} d_{13}.$$

But this is also equal to $\tilde{a}^t M \tilde{a}$ where

$$\tilde{a}^t = (|a_{11}|, |a_{22|}, |a_{33}|, \epsilon_2 a_{12}, \epsilon_3 a_{13})$$

which is in $P_{\mathcal{G}}$. Therefore $\langle K, D \rangle > 0$ for all $K \in P_{\mathcal{G}}$ if and only if $M$ is positive definite which translates immediately into the conditions defining $P_{\mathcal{G}}^*$ in Proposition 4.3.

**Proof of Theorem 4.4** For the proof of the theorem, it will be convenient to adopt a slightly different form of the parametrization of the Cholesky decomposition of $K = AA^t$ in $P_{\mathcal{G}}$. Let

$$A_{ij} = \begin{cases} \sqrt{a_{ii}} & \text{if } i = j, \\ -a_{ij} & \text{if } i < j. \end{cases}$$

so that

$$(AA^T)_{ij} = \begin{cases} a_{ii} + \sum_{l > i} a_{il}^2 & \text{if } i = j, \\ -a_{ij}\sqrt{a_{jj}} + \sum_{l > max(i,j)} a_{il} a_{jl} & \text{if } i < j. \end{cases}$$

Equating each entry $k_{ij}$ of $K$ to the corresponding entry of $AA^T$ with the constraint that $k_{13} = k_{23}$ shows that

$$k_{11} = a_{11} + a_{12}^2 + a_{13}^2, \quad k_{12} = -\sqrt{a_{22}} a_{12} + a_{13} a_{23},$$

30

$$k_{22} = a_{22} + a_{23}^2, \qquad k_{13} = -\sqrt{a_{33}}a_{13},$$

$$k_{33} = a_{33}, \qquad k_{23} = -\sqrt{a_{33}}a_{23}.$$

In particular, we find that since $a_{33} > 0$, $a_{13} = a_{23}$ and $k_{12} = -\sqrt{a_{22}}a_{12} + a_{13}^2$. The Jacobian of the transformation from $K$ to $A$ is

$$
J = \begin{array}{c} \\ a_{11} \\ a_{12} \\ a_{13} \\ a_{22} \\ a_{33} \end{array}
\begin{array}{ccccc} k_{11} & k_{12} & k_{13} & k_{22} & k_{33} \\ \left( \begin{array}{ccccc} 1 & 0 & 0 & 0 & 0 \\ * & -\sqrt{a_{22}} & 0 & 0 & 0 \\ * & * & -\sqrt{a_{33}} & 2a_{13} & 0 \\ * & * & * & 1 & 0 \\ * & * & * & * & 1 \end{array} \right) \end{array}
$$

It is easy to see $|J| = |diag(J)| = a_{22}^{1/2}a_{33}^{1/2}$.

We now have all the ingredients necessary to calculate the normalizing constant $I_{\mathcal{G}}(\delta, D)$. We have $|K| = a_{11}a_{22}a_{33}$ and

$$
\begin{aligned}
\langle K, D \rangle &= d_{11}k_{11} + d_{22}k_{22} + d_{33}k_{33} + 2d_{12}k_{12} + 2d_{13}k_{13} + 2d_{23}k_{23} \\
&= d_{11}(a_{11} + a_{12}^2 + a_{13}^2) + d_{22}(a_{22} + a_{23}^2) + d_{33}a_{33} \\
&\quad + 2d_{12}(-\sqrt{a_{22}}a_{12} + a_{13}a_{23}) + 2d_{13}(-a_{13}\sqrt{a_{33}}) + 2d_{23}(-a_{23}\sqrt{a_{33}}).
\end{aligned}
$$

and so the normalizing constant is

$$
\begin{aligned}
I_{\mathcal{G}}(\delta, D) &= \int_A a_{11}^{\frac{\delta-2}{2}} a_{22}^{\frac{\delta-1}{2}} a_{33}^{\frac{\delta-1}{2}} \exp(-\frac{1}{2}d_{11}a_{11} - \frac{1}{2}d_{22}a_{22} - \frac{1}{2}d_{33}a_{33} - \frac{1}{2}d_{11}a_{12}^2 \\
&\quad -\frac{1}{2}(d_{11} + d_{22} + 2d_{12})a_{13}^2 + d_{12}\sqrt{a_{22}}a_{12} + 2d_{13}a_{13}\sqrt{a_{33}})dA.
\end{aligned}
$$

where $a_{ii} > 0$; $a_{ij} \in R, i < j$; and $dA$ denotes the product of all differentials. The integral with respect to $a_{11}$ is a gamma integral with

$$
\int_0^\infty a_{11}^{\frac{\delta-2}{2}} \exp(-\frac{1}{2}d_{11}a_{11})da_{11} = 2^{\frac{\delta}{2}}\Gamma(\frac{\delta}{2})d_{11}^{-\frac{\delta}{2}}.
$$

The integral with respect to $a_{12}$ and $a_{13}$ are Gaussian integrals with

$$
\int_{-\infty}^\infty \exp(-\frac{1}{2}d_{11}a_{12}^2 + d_{12}\sqrt{a_{22}}a_{12})da_{12} = \frac{\sqrt{2\pi}}{\sqrt{d_{11}}} \exp(\frac{d_{12}^2 a_{22}}{2d_{11}}),
$$

31

and

$$\int_{-\infty}^{\infty} \exp(-\frac{1}{2}(d_{11} + d_{22} + 2d_{12})a_{13}^2 + 2d_{13}\sqrt{a_{33}}a_{13})da_{13} = \frac{\sqrt{2\pi}}{\sqrt{d_{11} + d_{22} + 2d_{12}}} \exp(\frac{2d_{13}^2 a_{33}}{d_{11} + d_{22} + 2d_{12}}).$$

Therefore

$$
\begin{aligned}
I_{\mathcal{G}}(\delta, D) &= \Gamma(\frac{\delta}{2})2^{\frac{\delta}{2}}d_{11}^{-\frac{\delta+1}{2}}2\pi(d_{11} + d_{22} + 2d_{12})^{-\frac{1}{2}} \\
&\quad \int_0^{\infty} a_{22}^{\frac{\delta-1}{2}} a_{33}^{\frac{\delta-1}{2}} \exp\{(-\frac{1}{2}d_{22} + \frac{d_{12}^2}{2d_{11}})a_{22} + (-\frac{1}{2}d_{33} + \frac{2d_{13}^2}{d_{11} + d_{22} + 2d_{12}})a_{33}\}da_{22}da_{33} \\
&= \Gamma(\frac{\delta}{2})2^{\frac{\delta}{2}}d_{11}^{-\frac{\delta+1}{2}}2\pi(d_{11} + d_{22} + 2d_{12})^{-\frac{1}{2}} \\
&\quad \times\Gamma(\frac{\delta+1}{2})(\frac{2d_{11}}{d_{11}d_{22} - d_{12}^2})^{\frac{\delta+1}{2}}\Gamma(\frac{\delta+1}{2})(\frac{2(d_{11} + d_{22} + 2d_{12})}{d_{33}(d_{11} + d_{22} + 2d_{12}) - 4d_{13}^2})^{\frac{\delta+1}{2}} \\
&= \Gamma(\frac{\delta}{2})\Gamma^2(\frac{\delta+1}{2})\pi 2^{\frac{3\delta+4}{2}}(d_{11} + d_{22} + 2d_{12})^{\frac{\delta}{2}}[d_{33}(d_{11} + d_{22} + 2d_{12}) - 4d_{13}^2]^{-\frac{\delta+1}{2}} \\
&\quad \times(d_{11}d_{22} - d_{12}^2)^{-\frac{\delta+1}{2}} .
\end{aligned}
$$

**Proof of Proposition 4.4**

We proceed as in the proof of Proposition 4.3. That is, we let $K = AA^t$ be the Cholesky decomposition of $K$ with $A$ upper triangular. Equating the entries of $K$ and $AA^t$ yields

$$a_{23} = a_{13}, \ a_{24} = a_{14}, \ a_{44} = a_{33}$$

with then

$$
\begin{aligned}
k_{11} &= a_{11}^2 + a_{12}^2 + a_{12}^2 + a_{14}^2, \quad k_{12} = a_{12}a_{22} + a_{13}^2 + a_{14}^2, \ k_{13} = a_{13}a_{33}. \quad k_{14} = a_{14}a_{33} \\
&\quad\quad\quad\quad\quad\quad\quad\quad k_{22} = a_{22}^2 + a_{13}^2 + a_{14}^2, \ k_{23} = a_{13}a_{33}, \quad k_{14} = a_{14}a_{33} \\
&\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad k_{33} = a_{33}^2 \quad\quad\quad\quad\quad\quad k_{34} = 0 \\
&\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad k_{44} = a_{33}^2
\end{aligned}
$$

Then, ordering $\langle K, D \rangle$ as a polynomial in $a_{ij}$, we see that

$$
\begin{aligned}
\langle K, D \rangle &= d_{11}a_{11}^2 + d_{22}a_{22}^2 + 2d_{33}a_{33}^2 + d_{11}a_{12}^2 + 2d_{12}a_{22}a_{12} + a_{13}^2(d_{11} + 2d_{12} + d_{22}) \\
&\quad + 4d_{13}a_{13}a_{33} + a_{14}^2(d_{11} + 2d_{12} + d_{22}) + 4d_{14}a_{14}a_{33}
\end{aligned}
$$

is a quadratic form and the matrix of this quadratic form is

$$
M = \begin{pmatrix}
d_{11} & 0 & 0 & 0 & 0 & 0 \\
0 & d_{22} & d_{12} & 0 & 0 & 0 \\
0 & d_{12} & d_{11} & 0 & 0 & 0 \\
0 & 0 & 0 & 2d_{33} & 2d_{13} & 2d_{14} \\
0 & 0 & 0 & 2d_{13} & d_{11} + 2d_{12} + d_{22} & 0 \\
0 & 0 & 0 & 2d_{14} & 0 & d_{11} + 2d_{12} + d_{22}
\end{pmatrix}.
$$

With exactly the same argument as in Proposition 4.3, we can show that $\langle K, D \rangle > 0$ for all $K \in \bar{P}_\mathcal{G}$ if and only if $M > 0$, i.e. $D$ satisfies the conditions of Proposition 4.4.

**Proof of Theorem 4.5** As in the proof of Theorem 4.4, it will be convenient to adopt a slightly different parametrization of the Cholesky decomposition of $K$. Let

$$
A_{ij} = \begin{cases}
\sqrt{a_{ii}} & \text{if } i = j, \\
-a_{ij} & \text{if } i < j.
\end{cases}
$$

so that the entries of $AA^t$ are given by

$$
(AA^T)_{ij} = \begin{cases}
a_{ii} + \sum\limits_{l>i} a_{il}^2 & \text{if } i = j, \\
-a_{ij}\sqrt{a_{jj}} + \sum\limits_{l>\max(i,j)} a_{il}a_{jl} & \text{if } i < j.
\end{cases}
$$

Equating each entry $k_{ij}$ of $K$ to the corresponding entry of $AA^T$, we find that

$$k_{11} = a_{11} + a_{12}^2 + a_{13}^2 + a_{14}^2, \qquad k_{12} = -\sqrt{a_{22}}a_{12} + a_{13}a_{23} + a_{14}a_{24},$$

$$k_{13} = -\sqrt{a_{33}}a_{13} + a_{14}a_{34}, \qquad k_{14} = -\sqrt{a_{44}}a_{14},$$

$$k_{22} = a_{22} + a_{23}^2 + a_{24}^2, \qquad k_{23} = -\sqrt{a_{33}}a_{23} + a_{24}a_{34},$$

$$k_{24} = -\sqrt{a_{44}}a_{24} \qquad k_{33} = a_{33} + a_{34}^2,$$

$$k_{34} = -\sqrt{a_{44}}a_{34}, \qquad k_{44} = a_{44}.$$

This shows that $a_{44} > 0$ and $a_{34} = 0$. Since $a_{33} > 0$ and $k_{13} = k_{23}$, then $a_{13} = a_{23}$. Since $a_{44} > 0$ and $k_{14} = k_{24}$, then $a_{14} = a_{24}$. Since $k_{34} = 0$, then $a_{33} = a_{44}$. Therefore, we obtain that

$$k_{11} = a_{11} + a_{12}^2 + a_{13}^2 + a_{14}^2, \qquad k_{12} = -\sqrt{a_{22}}a_{12} + a_{13}^2 + a_{14}^2,$$

$$k_{13} = k_{23} = -\sqrt{a_{33}}a_{13}, \qquad k_{14} = k_{24} = -\sqrt{a_{33}}a_{14},$$

$$k_{22} = a_{22} + a_{13}^2 + a_{14}^2, \qquad k_{33} = k_{44} = a_{33}.$$

The Jacobian of the transformation from $K$ to $A$ is

$$
J = \begin{array}{c} \\ a_{11} \\ a_{12} \\ a_{13} \\ a_{14} \\ a_{22} \\ a_{33} \end{array}
\begin{array}{cccccc}
k_{11} & k_{12} & k_{13} & k_{14} & k_{22} & k_{33} \\
\left( \begin{array}{cccccc}
1 & 0 & 0 & 0 & 0 & 0 \\
* & -\sqrt{a_{22}} & 0 & 0 & 0 & 0 \\
* & * & -\sqrt{a_{33}} & 0 & 2a_{13} & 0 \\
* & * & * & -\sqrt{a_{33}} & 2a_{14} & 0 \\
* & * & * & * & 1 & 0 \\
* & * & * & * & * & 1
\end{array} \right)
\end{array}
$$

It is easy to see $|J| = |diag(J)| = a_{22}^{1/2} a_{33}$. We now have all the ingredients necessary to calculate the normalizing constant $I_{G_2}(\delta, D)$. Through the change of variables, $K = AA^T$. Then $|K| = a_{11}a_{22}a_{33}^2$,

$$
\begin{aligned}
\langle K, D \rangle &= d_{11}k_{11} + d_{22}k_{22} + d_{33}k_{33} + d_{44}k_{44} + 2d_{12}k_{12} + 2d_{13}k_{13} + 2d_{14}k_{14} + 2d_{23}k_{23} + 2d_{24}k_{24} + 2d_{34}k_{34} \\
&= d_{11}(a_{11} + a_{12}^2 + a_{13}^2 + a_{14}^2) + d_{22}(a_{22} + a_{13}^2 + a_{14}^2) + 2d_{33}a_{33} \\
&\quad + 2d_{12}(-\sqrt{a_{22}}a_{12} + a_{13}^2 + a_{14}^2) + 4d_{13}(-a_{13}\sqrt{a_{33}}) + 4d_{14}(-a_{14}\sqrt{a_{33}}).
\end{aligned}
$$

and so the integral equals

$$
\begin{aligned}
I_\mathcal{G}(\delta, D) &= \int_A a_{11}^{\frac{\delta-2}{2}} a_{22}^{\frac{\delta-1}{2}} a_{33}^{\delta-1} \exp\{-\frac{1}{2}d_{11}a_{11} - \frac{1}{2}d_{11}a_{12}^2 + d_{12}\sqrt{a_{22}}a_{12} - \frac{1}{2}(d_{11} + d_{22} + 2d_{12})a_{13}^2 \\
&\quad + 2d_{13}a_{13}\sqrt{a_{33}} - \frac{1}{2}(d_{11} + d_{22} + 2d_{12})a_{14}^2 + 2d_{14}a_{14}\sqrt{a_{33}} - \frac{1}{2}d_{22}a_{22} - d_{33}a_{33}\}dA.
\end{aligned}
$$

where $a_{ii} > 0$; $a_{ij} \in R, i < j$; and $dA$ denotes the product of all differentials. The integral with respect to $a_{11}$ is gamma integrals, then

$$
\int_0^\infty a_{11}^{\frac{\delta-2}{2}} \exp(-\frac{1}{2}d_{11}a_{11})da_{11} = 2^{\frac{\delta}{2}}\Gamma(\frac{\delta}{2})d_{11}^{-\frac{\delta}{2}}.
$$

The integral with respect to $a_{12}$, $a_{13}$ and $a_{14}$ are normal integrals, then

$$
\int_{-\infty}^\infty \exp(-\frac{1}{2}d_{11}a_{12}^2 + d_{12}\sqrt{a_{22}}a_{12})da_{12} = \frac{\sqrt{2\pi}}{\sqrt{d_{11}}}\exp(\frac{d_{12}^2 a_{22}}{2d_{11}}),
$$

$$
\int_{-\infty}^\infty \exp(-\frac{1}{2}(d_{11} + d_{22} + 2d_{12})a_{13}^2 + 2d_{13}\sqrt{a_{33}}a_{13})da_{13} = \frac{\sqrt{2\pi}}{\sqrt{d_{11} + d_{22} + 2d_{12}}}\exp(\frac{2d_{13}^2 a_{33}}{d_{11} + d_{22} + 2d_{12}}),
$$

and

$$\int_{-\infty}^{\infty} \exp(-\frac{1}{2}(d_{11} + d_{22} + 2d_{12})a_{14}^2 + 2d_{14}\sqrt{a_{33}}a_{14})da_{14} = \frac{\sqrt{2\pi}}{\sqrt{d_{11} + d_{22} + 2d_{12}}} \exp(\frac{2d_{14}^2 a_{33}}{d_{11} + d_{22} + 2d_{12}}).$$

Therefore, the integral becomes

$$
\begin{aligned}
I_{G_1}(\delta, D) &= \Gamma(\frac{\delta}{2})2^{\frac{\delta}{2}}d_{11}^{-\frac{\delta+1}{2}}(2\pi)^{\frac{3}{2}}(d_{11} + d_{22} + 2d_{12})^{-1} \\
&\quad \int_0^{\infty} a_{22}^{\frac{\delta-1}{2}} a_{33}^{\frac{\delta-1}{2}} \exp\{(-\frac{1}{2}d_{22} + \frac{d_{12}^2}{2d_{11}})a_{22} + (-d_{33} + \frac{2d_{13}^2 + 2d_{14}^2}{d_{11} + d_{22} + 2d_{12}})a_{33}\}da_{22}da_{33} \\
&= \Gamma(\frac{\delta}{2})2^{\frac{\delta+3}{2}}d_{11}^{-\frac{\delta+1}{2}}\pi^{\frac{3}{2}}(d_{11} + d_{22} + 2d_{12})^{-1} \\
&\quad \Gamma(\frac{\delta+1}{2})(\frac{2d_{11}}{d_{11}d_{22} - d_{12}^2})^{\frac{\delta+1}{2}}\Gamma(\delta)(\frac{d_{11} + d_{22} + 2d_{12}}{d_{33}(d_{11} + d_{22} + 2d_{12}) - 2(d_{13}^2 + d_{14}^2)})^{\delta} \\
&= \Gamma(\frac{\delta}{2})\Gamma(\frac{\delta+1}{2})\Gamma(\delta)\pi^{\frac{3}{2}}2^{\delta+2}(d_{11} + d_{22} + 2d_{12})^{\delta-1}[d_{33}(d_{11} + d_{22} + 2d_{12}) - 2(d_{13}^2 + d_{14}^2)]^{-\delta} \\
&\quad (d_{11}d_{22} - d_{12}^2)^{-\frac{\delta+1}{2}}.
\end{aligned}
$$

# Appendix 2

## Numerical values for $D, E(K)$ and $\hat{K}$ and plots for Section 5

We give here the matrices $D$, $E(K)$ and $\hat{K}$ as well as the traceplot and autocorrelation plot of $\log(|K|)$ for all graphs in Fig. 1. Here $\hat{K}$ has been computed with 5000 iterations after a 1000 iterations burn in and averaged over 100 simulations

**Graph in Fig. 1(a)**

$$D = \begin{pmatrix} 1 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 5 & 2 & 0 & 0 & 0 \\ 2 & 2 & 2 & 25 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 6 & 2 & 2 \\ 0 & 0 & 0 & 0 & 2 & 3 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 4 \end{pmatrix},$$
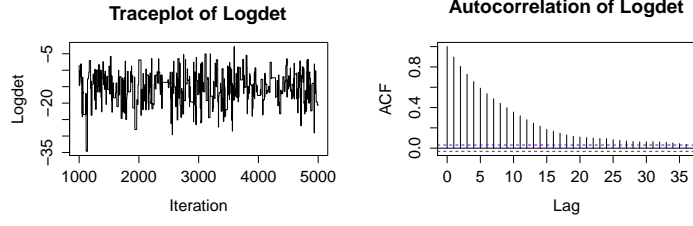
(a) Traceplot                    (b) ACF plot

Figure 3: (a) Traceplot of $\log(|K|)$ v.s. the number of iterations. (b) Autocorrelation plot of $\log(|K|)$ for Fig. 1(a).

$$
E(K) = \begin{pmatrix}
1.1294 & 0 & 0 & -0.0129 & 0 & 0 & 0 \\
0 & 0.5915 & 0 & -0.0129 & 0 & 0 & 0 \\
0 & 0 & 0.2578 & -0.0129 & 0 & 0 & 0 \\
-0.0129 & -0.0129 & -0.0129 & 0.0767 & -0.0129 & 0 & 0 \\
0 & 0 & 0 & -0.0129 & 0.2589 & -0.0129 & -0.0129 \\
0 & 0 & 0 & 0 & -0.0129 & 0.3699 & 0 \\
0 & 0 & 0 & 0 & -0.0129 & 0 & 0.2817
\end{pmatrix},
$$

$$
\hat{K} = \begin{pmatrix}
1.1274 & 0 & 0 & -0.0127 & 0 & 0 & 0 \\
0 & 0.5961 & 0 & -0.0127 & 0 & 0 & 0 \\
0 & 0 & 0.2563 & -0.0127 & 0 & 0 & 0 \\
-0.0127 & -0.0127 & -0.0127 & 0.0767 & -0.0127 & 0 & 0 \\
0 & 0 & 0 & -0.0127 & 0.2594 & -0.0127 & -0.0127 \\
0 & 0 & 0 & 0 & -0.0127 & 0.3708 & 0 \\
0 & 0 & 0 & 0 & -0.0127 & 0 & 0.2818
\end{pmatrix}.
$$

36

**Graph in Fig. 1(b)**

$$
D = \begin{pmatrix}
9 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
1 & 25 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
2 & 0 & 25 & 0 & 0 & 0 & 0 & 0 & 0 \\
3 & 0 & 0 & 25 & 0 & 0 & 0 & 0 & 0 \\
4 & 0 & 0 & 0 & 25 & 0 & 0 & 0 & 0 \\
5 & 0 & 0 & 0 & 0 & 25 & 0 & 0 & 0 \\
6 & 0 & 0 & 0 & 0 & 0 & 25 & 0 & 0 \\
7 & 0 & 0 & 0 & 0 & 0 & 0 & 25 & 0 \\
8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 25
\end{pmatrix}.
$$

$$
E(K) = \begin{pmatrix}
1.4778 & -0.0112 & -0.0225 & -0.0338 & -0.0451 & -0.0563 & -0.0676 & -0.0789 & -0.0902 \\
-0.0112 & 0.1015 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-0.0225 & 0 & 0.1015 & 0 & 0 & 0 & 0 & 0 & 0 \\
-0.0338 & 0 & 0 & 0.1015 & 0 & 0 & 0 & 0 & 0 \\
-0.0451 & 0 & 0 & 0 & 0.1015 & 0 & 0 & 0 & 0 \\
-0.0563 & 0 & 0 & 0 & 0 & 0.1015 & 0 & 0 & 0 \\
-0.0676 & 0 & 0 & 0 & 0 & 0 & 0.1015 & 0 & 0 \\
-0.0789 & 0 & 0 & 0 & 0 & 0 & 0 & 0.1015 & 0 \\
-0.0902 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.1015
\end{pmatrix}.
$$

$$
\hat{K} = \begin{pmatrix}
1.4690 & -0.0113 & -0.0223 & -0.0341 & -0.0455 & -0.0569 & -0.0677 & -0.0796 & -0.0905 \\
-0.0113 & 0.1016 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-0.0223 & 0 & 0.1016 & 0 & 0 & 0 & 0 & 0 & 0 \\
-0.0341 & 0 & 0 & 0.1016 & 0 & 0 & 0 & 0 & 0 \\
-0.0455 & 0 & 0 & 0 & 0.1016 & 0 & 0 & 0 & 0 \\
-0.0569 & 0 & 0 & 0 & 0 & 0.1016 & 0 & 0 & 0 \\
-0.0677 & 0 & 0 & 0 & 0 & 0 & 0.1016 & 0 & 0 \\
-0.0796 & 0 & 0 & 0 & 0 & 0 & 0 & -0.1016 & 0 \\
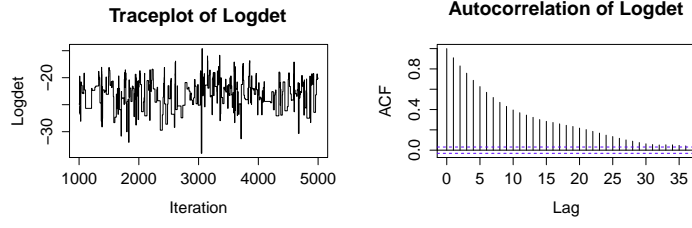-0.0905 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.1016
\end{pmatrix}.
$$

**Traceplot of Logdet**

**Autocorrelation of Logdet**
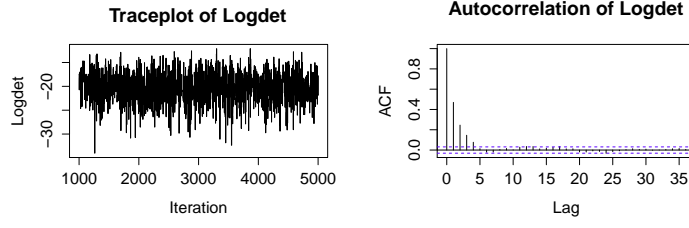
(a) Traceplot

(b) ACF plot

Figure 4: Traceplot and Autocorrelation plot of $\log(|K|)$ for Graph in Fig. 1(b).

**Graph in Fig. 1(c)**

$$
D = \begin{pmatrix}
25 & 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\
9 & 25 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
8 & 0 & 25 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
7 & 0 & 0 & 25 & 0 & 0 & 0 & 0 & 0 & 0 \\
6 & 0 & 0 & 0 & 25 & 0 & 0 & 0 & 0 & 0 \\
5 & 0 & 0 & 0 & 0 & 25 & 0 & 0 & 0 & 0 \\
4 & 0 & 0 & 0 & 0 & 0 & 25 & 0 & 0 & 0 \\
3 & 0 & 0 & 0 & 0 & 0 & 0 & 25 & 0 & 0 \\
2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 25 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 25
\end{pmatrix} .
$$

$$
E(K) = \begin{pmatrix}
0.1229 & -0.0013 & -0.0026 & -0.0039 & -0.0052 & -0.0065 & -0.0078 & -0.0091 & -0.0104 & -0.0117 \\
-0.0013 & 0.1229 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-0.0026 & 0 & 0.1229 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-0.0039 & 0 & 0 & 0.1229 & 0 & 0 & 0 & 0 & 0 & 0 \\
-0.0052 & 0 & 0 & 0 & 0.1229 & 0 & 0 & 0 & 0 & 0 \\
-0.0065 & 0 & 0 & 0 & 0 & 0.1229 & 0 & 0 & 0 & 0 \\
-0.0078 & 0 & 0 & 0 & 0 & 0 & 0.1229 & 0 & 0 & 0 \\
-0.0091 & 0 & 0 & 0 & 0 & 0 & 0 & 0.1229 & 0 & 0 \\
-0.0104 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.1229 & 0 \\
-0.0117 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.1229
\end{pmatrix} .
$$

(a) Traceplot          (b) ACF plot

Figure 5: Traceplot and Autocorrelation plot of $\log(|K|)$ for Graph in Fig. 1(c).

$$\hat{K} = \begin{pmatrix} 0.1223 & -0.0012 & -0.0027 & -0.0041 & -0.0055 & -0.0064 & -0.0077 & -0.0090 & -0.0102 & -0.0115 \\ -0.0012 & 0.1223 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.0027 & 0 & 0.1223 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.0041 & 0 & 0 & 0.1223 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.0055 & 0 & 0 & 0 & 0.1223 & 0 & 0 & 0 & 0 & 0 \\ -0.0064 & 0 & 0 & 0 & 0 & 0.1223 & 0 & 0 & 0 & 0 \\ -0.0077 & 0 & 0 & 0 & 0 & 0 & 0.1223 & 0 & 0 & 0 \\ -0.0090 & 0 & 0 & 0 & 0 & 0 & 0 & 0.1223 & 0 & 0 \\ -0.0102 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.1223 & 0 \\ -0.0115 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.1223 \end{pmatrix}.$$

**Graph in Fig. 1(d)**

$$D = \begin{pmatrix} 3 & 1 & 2 \\ 1 & 4 & 2 \\ 2 & 2 & 5 \end{pmatrix}, \quad E(K) = \begin{pmatrix} 1.8108 & -0.0073 & -0.5517 \\ -0.0073 & 1.4472 & -0.5517 \\ -0.5517 & -0.5517 & 1.2413 \end{pmatrix}$$
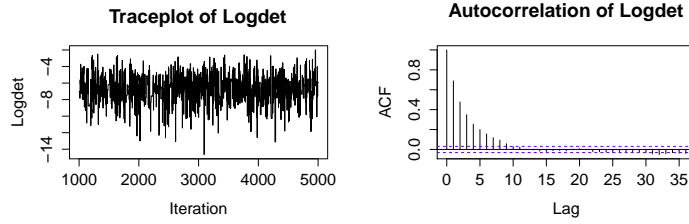
and

$$\hat{K} = \begin{pmatrix} 1.8097 & -0.0075 & -0.5514 \\ -0.0075 & 1.4485 & -0.5514 \\ -0.5514 & -0.5514 & 1.2442 \end{pmatrix}.$$
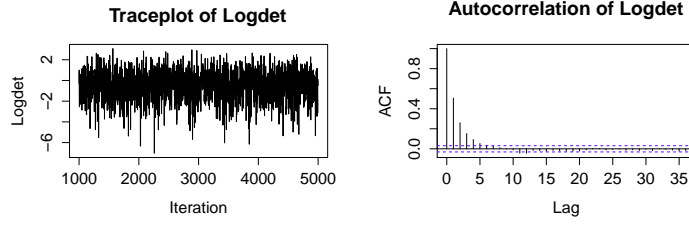
(a) Traceplot

(b) ACF plot

Figure 6: Traceplot and Autocorrelation plot of $\log(|K|)$ for Graph in Fig. 1(d).



(a) Traceplot

(b) ACF plot

Figure 7: Traceplot and Autocorrelation plot of $\log(|K|)$ for Graph in Fig. 1(e).

**Graph in Fig. 1 (e)**

$$
D = \begin{pmatrix} 2 & 1 & 3 & 4 \\ 1 & 1 & 3 & 4 \\ 3 & 3 & 200 & 0 \\ 4 & 4 & 0 & 200 \end{pmatrix}, \quad E(K) = \begin{pmatrix} 4.4631 & -3.5368 & -0.0189 & -0.0252 \\ -3.5368 & 8.4631 & -0.0189 & -0.0252 \\ -0.0189 & -0.0189 & 0.0157 & 0 \\ -0.0252 & -0.0252 & 0 & 0.0157 \end{pmatrix}
$$

and

$$
\hat{K} = \begin{pmatrix} 4.4714 & -3.5386 & -0.0192 & -0.0256 \\ -3.5386 & 8.4658 & -0.0192 & -0.0256 \\ -0.0192 & -0.0192 & 0.0158 & 0 \\ -0.0256 & -0.0256 & 0 & 0.0158 \end{pmatrix}.
$$

40

# Appendix 3

**Estimates and batch standard errors for entries of $K$ for the models of Section 6**

The estimates and batch standard errors are given below for the entries of $K$ listed in lexicographic order.

Table 4: The average estimates for entries of $K$ for Fig. 2 (b) when $p = 20$

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 0.1072 | 0.0100 | 0.0096 | 0.0322 | 0.0109 | 0.0093 | 0.0102 | 0.0105 | 0.0101 |
| 0.0103 | 0.0099 | 0.0106 | 0.0100 | 0.0099 | 0.0109 | 0.0104 | 0.0111 | 0.0116 |
| 0.0104 | 0.0100 | 0.0113 | 0.0115 | | | | | |

Table 5: The batch standard errors for Fig. 2 (b) when $p = 20$

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 0.0004 | 0.0005 | 0.0005 | 0.0001 | 0.0005 | 0.0005 | 0.0005 | 0.0005 |
| 0.0005 | 0.0005 | 0.0005 | 0.0004 | 0.0005 | 0.0004 | 0.0005 | 0.0004 |
| 0.0005 | 0.0005 | 0.0005 | 0.0005 | 0.0004 | 0.0005 | | |

Table 6: The average estimates for entries of $K$ for Fig. 2 (b) when $p = 30$

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 0.1217 | 0.0109 | 0.0126 | 0.0366 | 0.0109 | 0.0118 | 0.0120 | 0.0120 | 0.0115 | 0.0122 |
| 0.0108 | 0.0121 | 0.0113 | 0.0119 | 0.0125 | 0.0114 | 0.0120 | 0.0112 | 0.0119 | 0.0131 |
| 0.0115 | 0.0125 | 0.0116 | 0.0132 | 0.0110 | 0.0119 | 0.0119 | 0.0107 | 0.0129 | 0.0119 |
| 0.0124 | 0.0119 | | | | | | | | |

Table 7: The batch standard errors for Fig. 2 (b) when $p = 30$

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 0.0008 | 0.0006 | 0.0006 | 0.0003 | 0.0006 | 0.0006 | 0.0006 | 0.0006 |
| 0.0006 | 0.0006 | 0.0006 | 0.0006 | 0.0006 | 0.0006 | 0.0006 | 0.0006 |
| 0.0006 | 0.0006 | 0.0006 | 0.0006 | 0.0006 | 0.0006 | 0.0006 | 0.0006 |
| 0.0006 | 0.0005 | 0.0006 | 0.0005 | 0.0006 | 0.0006 | 0.0005 | 0.0005 |

Table 8: The average estimates for entries of $K$ for Fig. 2 (c) when $p = 20$

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 0.1102 | 0.0106 | 0.0104 | 0.0347 | 0.1135 | 0.0329 | 0.1104 | 0.0335 | 0.1113 | 0.0332 |
| 0.1103 | 0.0326 | 0.1157 | 0.0330 | 0.1082 | 0.0333 | 0.1083 | 0.0318 | 0.1096 | 0.0326 |
| 0.1059 | 0.0311 | | | | | | | | |

Table 9: The batch standard errors for Fig. 2 (c) when $p = 20$

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 0.0011 | 0.0002 | 0.0002 | 0.0004 | 0.0012 | 0.0003 | 0.0011 | 0.0004 |
| 0.0012 | 0.0004 | 0.0012 | 0.0003 | 0.0013 | 0.0003 | 0.0012 | 0.0003 |
| 0.0012 | 0.0004 | 0.0011 | 0.0003 | 0.0012 | 0.0003 | | |

Table 10: The average estimates for entries of $K$ for Fig. 2 (c) when $p = 30$

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 0.1295 | 0.0117 | 0.0111 | 0.0384 | 0.1253 | 0.0386 | 0.1266 | 0.0376 | 0.1248 | 0.0357 |
| 0.1214 | 0.0358 | 0.1209 | 0.0357 | 0.1181 | 0.0358 | 0.1161 | 0.0349 | 0.1126 | 0.0345 |
| 0.1123 | 0.0339 | 0.1126 | 0.0338 | 0.1136 | 0.0330 | 0.1143 | 0.0323 | 0.1083 | 0.0324 |
| 0.1077 | 0.0318 | | | | | | | | |

Table 11: The batch standard errors for Fig. 2 (c) when $p = 30$

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 0.0013 | 0.0002 | 0.0002 | 0.0003 | 0.0012 | 0.0004 | 0.0011 | 0.0004 |
| 0.0011 | 0.0003 | 0.0011 | 0.0003 | 0.0013 | 0.0004 | 0.0010 | 0.0004 |
| 0.0011 | 0.0003 | 0.0010 | 0.0003 | 0.0010 | 0.0003 | 0.0011 | 0.0003 |
| 0.0010 | 0.0003 | 0.0012 | 0.0003 | 0.0010 | 0.0003 | 0.0010 | 0.0003 |

# References

HØJSGAARD, S. & LAURITZEN, S. L. (2008). *Graphical Gaussian models with edge and vertex symmetries. J. R. Stat. Soc. Ser.* B, **70**, 1005-1027.

DOBRA, A., LENKOSKI, A. & RODRIGUEZ, A. (2011). *Bayesian inference for general Gaussian graphical models with application to multivariate lattice data. J. Am. Statist. Assoc.,* **106**, 1418–1433.

ATAY-KAYIS, A. & MASSAM, H. (2005). *A Monte Carlo method for computing the marginal likelihood in nondecomposable Gaussian graphical models. Biometrika*, **92**, 317-335.

DIACONIS, P. & Ylvisaker, D. (1979). *Conjugate priors for exponential families. Ann. Statist.*, **7**, 269-281.

LAURITZEN, S. L. (1996). *Graphical Models.* Oxford University Press.

LENKOSKI, A. (2013). *A direct sampler for G-Wishart variates. Stat*, **2**, 119-128.

WANG, H. & LI, S. Z. (2012). *Efficient Gaussian graphical model determination under G-Wishart distributions. Electron. J. Stat.*, **6**, 168-198.

PICCIONI, M. (2000). *Independence structure of natural conjugate densities to exponential families and the Gibbs Sampler. Scand. J. Statist.*, **27**, 111-127.

MENGERSEN, K. L. & TWEEDIE, R. L. (2012). *Rates of convergence of the Hastings and Metropolis algorithm. Ann. Statist.*, **24**, 101-121.

MITSAKAKIS, N., MASSAM, H. & ESCOBAR, M. D. (2011) *A Metropolis Hastings based method for sampling from the G-Wishart Distribution in Gaussian Graphical Models. Elect. J. of Statistics.*, **5**, 18-30.

PAULSEN, I., POWER, S. C., & SMITH, R. (1989) *Schur products and matrix completions. J. Funct. Anal.*, **85**, 151-78.

ROVERATO, A. (2000). *Cholesky Decomposition of a Hyper Inverse Wishart Matrix. Biometrika*, **87**, 99-112.

ROVERATO, A. (2002). *Hyper inverse Wishart distribution for non-decomposable graphs and its application to Bayesian inference for Gaussian graphical models. Scand. J. Statist.*, **29**, 391-411.

UHLER, C., LENKOSKI, A. & RICHARDS, D. (2014). *Exact formulas for the normalizing constant of the Wishart distributions for graphical models. http://arxiv.org/abs/1406.4901.*

WATSON, G.N.(1995). *A treatise on the theory of Bessel functions.* Cambridge University Press.